

UNIVERSIDADE DE LISBOA

Faculdade de Medicina de Lisboa



**Análise de STRs do cromossoma Y numa população do Sul de
Portugal com o *PowerPlex® Y23***

Marisa Isabel Ferreira Cruz

Orientador: Doutora Mónica Carvalho

Co-Orientador: Professor Doutor Jorge Costa Santos

Dissertação especialmente elaborada para a obtenção do grau de Mestre em
Medicina Legal e Ciências Forenses.

2016

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 19 de janeiro de 2016.

Agradecimentos

À minha orientadora interna do Serviço de Genética e Biologia Forenses, Doutora Mónica Carvalho, um primeiro e especial agradecimento por todos os ensinamentos, disponibilidade, paciência e companheirismo.

Ao Senhor Professor Doutor Jorge Costa Santos, simultaneamente, meu coorientador e, à altura diretor da Delegação Sul do Instituto Nacional de Medicina Legal e Ciências Forenses, um agradecimento também, especial, por permitir a realização do projeto nas instalações do Serviço de Genética e Biologia Forenses do Instituto Nacional de Medicina Legal e Ciências Forenses, I.P.

À Diretora do Serviço de Genética e Biologia Forenses do Instituto Nacional de Medicina Legal e Ciências Forenses, I.P., Doutora Maria João Porto, por ter autorizado a realização da tese no serviço que superiormente dirige.

À Dr.^a Teresa Ribeiro, responsável pelo Serviço de Genética e Biologia Forenses na Delegação Sul, pelo acolhimento e por facilitar a boa realização e concretização dos trabalhos.

Aos restantes elementos do Serviço de Genética e Biologia Forenses, nomeadamente ao Doutor António Amorim, à Mestre Heloísa Afonso Costa e à Mestre Cláudia Vieira da Silva uma palavra de reconhecimento e agradecimento pelo acolhimento e pelos dias de trabalho que comigo partilharam.

À Rita Oliveira Dario por todo o apoio a nível laboratorial e por todos os conselhos nos momentos mais difíceis.

À Dr.^a Manuela Marques pela ajuda na pesquisa bibliográfica, simpatia e disponibilidade demonstrada.

Às minhas colegas de Mestrado, Ana Resende, Ana Rita Inácio e Sara Gaudêncio por todos os momentos que passámos juntas, tornaram, sem dúvida, todo este percurso mais fácil.

Aos meus pais, José Cruz e Maria Clara Ferreira, por estarem sempre ao meu lado, pelo esforço feito e apoio incondicional, sem os quais não teria chegado onde cheguei. Um grande Obrigado.

Resumo

A Genética Forense apresenta um papel fundamental na resolução de perícias relacionadas com investigações de filiação, identificação genética individual e criminalística biológica. A partir de um vestígio biológico, é possível identificar um perfil genético que será comparado com o perfil determinado a partir de uma amostra biológica de referência, obtendo-se assim a prova genética.

Os STRs são particularmente abundantes no genoma humano e apresentam um elevado grau de polimorfismo, características que os tornaram nos marcadores mais usados em Genética Forense, apresentando um elevado poder de discriminação entre indivíduos de uma população.

Os marcadores genéticos autossómicos são os mais frequentemente utilizados para fins de identificação humana, no entanto existem situações em que estes, por si só, não permitem obter resultados conclusivos. É o caso da agressão sexual, onde poderá existir uma mistura de material biológico de origem feminina e masculina, estando o componente feminino presente em maior proporção. Ao analisar os marcadores autossómicos, pode não ser possível a deteção do perfil masculino, devido à desproporção dos componentes, o que não acontece quando se recorre ao estudo dos STRs do cromossoma Y. No entanto, há que ter em atenção o facto de a informação fornecida por estes marcadores ser partilhada por familiares com a mesma linhagem paterna, o que condiciona o grau de discriminação.

O *kit PowerPlex® Y23* possibilita a tipagem de 23 Y-STRs em simultâneo, sendo que a análise de 6 deles só é, presentemente, possível com este *multiplex*.

Estes novos *loci* apresentam um poder de discriminação superior em relação aos que compõem outros *kits* comerciais para análise de Y-STRs, possibilitando uma melhor distinção entre indivíduos não aparentados do sexo masculino.

Neste trabalho foram tipadas 243 amostras de indivíduos do sexo masculino, não aparentados residentes na região Sul de Portugal, com o *kit PowerPlex® Y23*. Observaram-se 237 haplótipos diferentes, dos quais 231 são únicos e 6 haplótipos são partilhados por dois indivíduos. Os resultados do cálculo dos respetivos parâmetros estatísticos forenses em conjunto com os estudos de validação do método, nomeadamente estudo de concordância e de artefactos da PCR, demonstram que os 23 marcadores genéticos são adequados para a utilização no âmbito de perícias de genética e biologia forense.

Palavras-chave: Y-STR, *kit PowerPlex® Y23*, Genética Populacional.

Abstract

Forensic genetics plays a fundamental role in the resolution of routine forensic casework related to kinship investigation, individual genetic identification and crime casework. From a biological sample it's possible to identify a genetic profile which will be compared with the one determined from a biological reference sample, therefore, obtaining a genetic proof.

STRs are particularly abundant in human genome and show a high degree of polymorphism, characteristics that have made them the most used markers in Forensic Genetics, exhibiting a significant discrimination power between individuals of a certain population.

Autosomal DNA markers are frequently used in human identification testing, however there are cases where these alone do not provide conclusive results. For example, in sexual assault cases where there may be a mixture of biological material of female and male origin, being the female component in higher proportion. The analysis of autosomal markers, may not allow the detection of a male profile, due to the disproportion of the components, which does not occur when using Y chromosome STRs. However, it should be noted the fact that, the information provided by these markers can be shared by relatives with the same paternal lineage, which affects the discrimination degree.

The PowerPlex® Y23 kit allows simultaneous genotyping of 23 Y-STR loci, and the analysis of 6 of them is, currently, only possible with this multiplex. These new loci exhibit a power of discrimination superior to those incorporated in other

commercial kits for Y-STRs analysis, providing an increased capacity for differentiating unrelated male individuals.

In this study, samples of 243 unrelated male individuals living in the south of Portugal were typed, with the PowerPlex® Y23 kit. A total of 237 different haplotypes were found, among which 231 were unique and 6 haplotypes were shared by two individuals. The results from the statistical parameters calculation, together with the method validation studies, namely concordance studies and PCR artefacts, prove that the 23 genetic markers are suitable for forensic applications.

Keywords: PowerPlex® Y23 kit, Y-STR, Population Genetics.

Índice

Agradecimentos.....	iii
Resumo	v
Abstract	vii
Índice.....	ix
Índice de figuras	xii
Índice de tabelas	xiii
Abreviaturas	xiv
I – Introdução.....	1
1.1. Genoma Humano – Estrutura e organização do ADN.....	2
1.2. Cromossoma Y.....	4
1.3. Marcadores Genéticos	7
1.3.1. Estrutura e Nomenclatura dos marcadores STRs.....	7
1.3.2. Polimorfismos do Cromossoma Y.....	9
1.3.3. Taxas de Mutação em Y-STRs	11
1.3.4. Kits Comerciais de marcadores genéticos.....	13
1.3.4.1. Sistema <i>AmpFISTR® Yfiler®</i>	14
1.3.4.2. Sistema <i>PowerPlex® Y23</i>	16
1.4. Genética populacional e parâmetros estatísticos de interesse forense.....	18
1.5. Aplicação Forense dos Y-STRs	20

1.6. Validação de métodos	22
1.7. Objetivos do estudo	26
II - Materiais e Métodos	27
2.1. Amostragem	27
2.2. Procedimento Experimental	27
2.2.1. Extração de ADN	28
2.2.2. Amplificação por PCR	29
2.2.3. Aplicação e Detecção dos produtos de PCR	32
2.2.4. Análise de resultados.....	34
2.3. Validação de Procedimento de Ensaio.....	34
2.3.1. Estudo de concordância	34
2.3.1.1. Comparação com <i>AmpFISTR</i> [®] <i>Yfiler</i> [®]	35
2.3.1.2. Volume de reação	36
2.3.2. Artefactos da PCR	36
2.3.3. Balanceamento dos heterozigóticos (DYS385a/b).....	37
III – Resultados e Discussão	39
3.1. Haplótipos do cromossoma Y.....	39
3.2. Análise Estatística	39
3.3. Variantes alélicas	47
3.4. Validação do método de ensaio	49
3.4.1. Estudo de concordância	49

3.4.1.1. Comparação com <i>AmpFISTR® Yfiler®</i>	49
3.4.1.2. Volume de reação	49
3.4.2. Artefactos da PCR	52
3.4.3. Balanceamento dos heterozigóticos (DYS385a/b).....	54
IV – Conclusões	56
V - Referências Bibliográficas.....	58
VI – Anexos	66
Anexo 1 – Haplótipos de STRs do cromossoma Y na população Sul de Portugal....	66
Anexo 2 – Frequências alélicas dos marcadores genéticos do <i>kit PowerPlex® Y23</i>	83
Anexo 3 - Frequência genotípica do marcador <i>DYS385a/b</i>	87
Anexo 4 – Artigo publicado	88

Índice de figuras

Figura 1 - Estrutura do cromossoma Y humano (Adaptado de Shamsi, Kumar e Dada, 2011).....	6
Figura 2 - Estrutura de um STR (Adaptada de Butler, 2010).....	8
Figura 3 - Fluxograma representativo das fases do procedimento experimental.	28
Figura 4 - Número de alelos detetados em cada marcador genético estudado pelo <i>kit PowerPlex® Y23</i>	40
Figura 5 - Distribuição alélica de cada marcador analisado com o <i>kit PowerPlex® Y23</i> , nos 243 indivíduos estudados.....	41
Figura 6 - Frequência genotípica do marcador DYS385a/b.....	44
Figura 7 - Eletroforegramas da mesma amostra amplificada com o <i>kit PowerPlex® Y23</i> com um volume final de 25µL (A), 12.5µL (B) e 6.25µL (C).	51
Figura 8 - Eletroforegrama exemplificativo do marcador DYS385a/b em equilíbrio alélico (A) e em desequilíbrio alélico (B).	54

Índice de tabelas

Tabela 1 - Marcadores incluídos no <i>kit AmpFISTR® Yfiler®</i> e respectivos alelos, estrutura repetitiva e taxa de mutação (Obtido no <i>site</i> http://yhrd.org/ , visualizado em Maio de 2015).....	15
Tabela 2 - Marcadores incluídos no <i>kit PowerPlex® Y23</i> e respectivos alelos, estrutura repetitiva e taxa de mutação (Obtido no <i>site</i> http://yhrd.org/ , visualizado em Maio de 2015).	17
Tabela 3 - Reagentes e volumes utilizados por amostra na preparação da mistura de reação para amplificação com o <i>kit PowerPlex® Y23</i>	31
Tabela 4 - Condições da reação de amplificação com o <i>kit PowerPlex® Y23</i> num termociclador <i>GeneAmp 9700</i>	31
Tabela 5 - Reagentes e volumes utilizados por amostra na preparação da mistura de eletroforese capilar com o <i>kit PowerPlex® Y23</i>	33
Tabela 6 - Reagentes e volumes utilizados por amostra na preparação da mistura de reação para amplificação com o <i>kit AmpFISTR® YFiler®</i>	35
Tabela 7 - Condições da reação de amplificação com o <i>kit PowerPlex® Y23</i> num termociclador <i>GeneAmp 9700</i>	36
Tabela 8 - Diversidade génica.	46
Tabela 9 - Microvariantes alélicas observadas.	48
Tabela 10 - Média dos valores de <i>stutter</i> para n-1 e n+1 por marcador, determinados nas 243 amostras analisadas com o <i>kit PowerPlex® Y23</i>	53

Abreviaturas

°C – Grau Celsius

µL – MicroLitro

A – Adenina

ADN – Ácido desoxirribonucleico

ADNmt – ADN mitocondrial

C – Citosina

ddNTPs – Didesoxinucleótidos trifosfatados

FSS – *Forensic Science Service*

g - Grama

G – Guanina

GD – Diversidade Génica (*Genetic Diversity*)

HD – Diversidade Haplotípica (*Haplotype Diversity*)

INMLCF – Instituto Nacional de Medicina Legal e Ciências Forenses

ISFG – *International Society for Forensic Genetics*

Mb – milhões de pares de bases

mL – Mililitro

NRY – Parte não recombinante do cromossoma Y (*Nonrecombining part of the Y-chromosome*)

PAR – Região pseudoautossómica (*Pseudoautosomal region*)

Pb – Pares de bases

PCR – Reação em cadeia da polimerase (Polymerase Chain Reaction)

rpm – Rotações por minuto

SGBF–DS – Serviço de Genética e Biologia Forenses, Delegação do Sul

SNPs – Polimorfismo de nucleótido único (*Single Nucleotide Polymorphism*)

SRY – *Sex-determining region Y*

STRs – Pequenas Repetições em *Tandem* (*Short Tandem Repeats*)

T - Timina

Y-STRs – *Short Tandem Repeats* do cromossoma Y

I – Introdução

A Genética Forense integra uma das muitas áreas da Medicina Legal, apresentando um papel fundamental na resolução de perícias relacionadas com investigações de filiação, identificação genética individual e criminalística biológica (Pinheiro, 2010). O grande desenvolvimento desta ciência remonta a meados da década de 1980, por Alec Jeffreys, na sequência da descoberta de determinadas regiões no genoma humano que se repetiam várias vezes, repetições essas variáveis em número e de indivíduo para indivíduo, e que viriam a possibilitar a realização de testes de identificação humana (Butler, 2010).

A grande maioria das amostras biológicas analisadas em laboratórios forenses é de origem humana. O protocolo a seguir para o seu tratamento irá sofrer alterações consoante estas sejam amostras de referência (amostras colhidas a indivíduos identificados) ou amostras problema (amostras que pela qualidade e/ou quantidade de ADN dificultam a obtenção de um perfil genético completo). Independentemente da origem destas, o seu manuseamento é feito com recurso a procedimentos que visem certificar a cadeia de custódia, a autenticidade e integridade das amostras (Pinheiro, 2013).

O objetivo final da perícia em Genética Forense será a obtenção de um perfil genético que se caracteriza por um conjunto de marcadores genéticos, que permitem a individualização e comparação com perfis previamente obtidos a partir de uma amostra biológica de referência, obtendo-se assim a prova genética, quer no âmbito criminal, quer civil (Pinheiro, 2010).

Após todo o processo laboratorial, adaptado a cada caso em particular, será elaborado um relatório pericial cujo conteúdo dependerá do tipo de investigação em causa. Caso se trate de uma investigação de parentesco biológico poderá haver uma exclusão de paternidade, sendo referidos os sistemas que conduziram a tal conclusão, ou não exclusão de paternidade, sendo estabelecida a probabilidade e o índice de paternidade. No campo da criminalística biológica, é feita a comparação dos perfis genéticos das amostras problema com as de referência, seguida da valoração estatística dos resultados no caso de uma coincidência de perfis ou compatibilidade (Butler, 2010; Pinheiro, 2010).

1.1. Genoma Humano – Estrutura e organização do ADN

A molécula de Ácido Desoxirribonucleico (ADN) é composta por duas cadeias polinucleotídicas constituídas por quatro tipos de nucleótidos. Cada nucleótido é composto por um açúcar desoxirribose, um grupo fosfato e uma base azotada que poderá ser uma adenina (A), citosina (C), guanina (G) ou timina (T). Estas podem ser classificadas em purinas (adenina e guanina) e em pirimidinas (timina e citosina). As duas cadeias polinucleotídicas unem-se através de pontes de hidrogénio estabelecidas entre as bases dos nucleótidos, mais especificamente entre A e T e entre G e C, existindo por isso uma complementaridade de bases que permite uma disposição energeticamente favorável no interior da dupla hélice (Alberts *et al.*, 2010).

A informação genética encontra-se codificada no ADN através de uma sequência de nucleótidos, e o conjunto de toda esta informação existente numa célula denomina-se genoma. Praticamente todo o ADN de uma célula encontra-se armazenado no seu núcleo, designando-se por isso ADN nuclear. Este encontra-se em cromossomas, cada um composto por uma única molécula de ADN linear, de cadeia dupla, com proteínas associadas que lhe conferem proteção (histonas), complexo que se denomina cromatina. O genoma humano é constituído por 22 pares de cromossomas autossómicos e 2 cromossomas sexuais. Cada célula humana contém um total de 46 cromossomas, ou seja, duas cópias de cada cromossoma, uma herdada da mãe e outra do pai, designados por cromossomas homólogos. O único par de cromossomas não-homólogos é o dos cromossomas sexuais, que no caso de um indivíduo do sexo masculino é herdado do pai um cromossoma Y e da mãe um cromossoma X. Caso se trate de um indivíduo do sexo feminino é herdado um cromossoma X de cada um dos progenitores (Goodwin, Linacre e Hadi, 2007).

É nos cromossomas que se localizam os genes, regiões codificantes do ADN que regulam, na maioria dos casos, a síntese de proteínas. São constituídos por sequências alternadas de exões, as sequências codificantes, e intrões, as sequências intercalares não codificantes. Como seria de esperar, existe uma correlação entre a complexidade de um organismo e o número de genes presentes no seu genoma. Contudo, os genes representam apenas cerca 2% (U.S. Department of Energy Human Genome Project, 2013) do genoma humano, o restante ADN é composto por regiões não codificantes por não estarem diretamente

envolvidas na síntese proteica, sendo porém essenciais para uma correta expressão génica de determinados genes (Alberts *et al.*, 2010; Butler, 2010).

Uma pequena percentagem do ADN celular total (1%) é composta por ADN mitocondrial (ADNmt), localizado nas mitocôndrias. Caracteriza-se por ser uma molécula circular, de cadeia dupla, com aproximadamente 16,569 pares de bases (pb). A cadeia dupla apresenta uma cadeia pesada (H) rica em purinas e uma cadeia leve (L) rica em pirimidinas. O ADN mitocondrial distingue-se pelo seu elevado número de cópias por célula, transmissão unicamente pela via materna, sem recombinação e elevada taxa de mutação, características que o tornam útil em estudos de evolução humana e na prática forense no caso de amostras com reduzida quantidade de ADN ou de ADN degradado (Primorac, Schanfield e Marjanović, 2014).

1.2. Cromossoma Y

A descoberta do cromossoma Y humano remonta a 1921 (Hughes e Rozen, 2012), tendo divergido de um ancestral autossómico (Bachtrog, 2014). Destaca-se de todos os outros cromossomas humanos por apenas se encontrar presente nas células de indivíduos do sexo masculino e ser transmitido unicamente de pai para filho (Noordam, 2012). É o segundo cromossoma mais pequeno do genoma humano, representando cerca de 1,8% deste, com uma média de 60 milhões de pares de bases (Mb) (Kayser e Ballantyne, 2014). Nas extremidades distais do braço curto e longo do cromossoma Y localizam-se duas regiões pseudoautossómicas

(PAR1 e PAR2), homólogas com regiões do cromossoma X, ocorrendo recombinação entre as regiões sub-teloméricas dos dois cromossomas (Figura 1). Assim, cerca de 95% do cromossoma Y não sofre recombinação durante a meiose, região esta denominada *nonrecombining part of the Y-chromosome* (NRY) (Kayser e Ballantyne, 2014), mantendo-se intacta ao longo de uma linhagem paterna, a menos que ocorra uma mutação (Gusmão *et al.*, 1999).

O cromossoma Y codifica 52 genes diferentes e ativos, igualmente distribuídos entre as regiões PAR (27 genes) e NRY (25 genes). Os genes presentes nas regiões PAR são herdados de forma semelhante aos genes dos cromossomas autossómicos, estando 24 deles localizados na região PAR1 e os restantes 3 na região PAR2 (Navarro-Costa, 2012). A região NRY, formada por heterocromatina e eucromatina (Figura 1), é rica em genes muitos dos quais estão implicados em importantes funções biológicas, como produção de esperma (Hughes e Rozen, 2012). É na região eucromática que se encontra um dos genes mais estudados, o gene SRY (situado no braço curto do cromossoma Y, Yp) determinante do sexo, que codifica uma proteína desencadeadora do desenvolvimento testicular (Goodwin, Linacre e Hadi, 2007; Noordam, 2012)

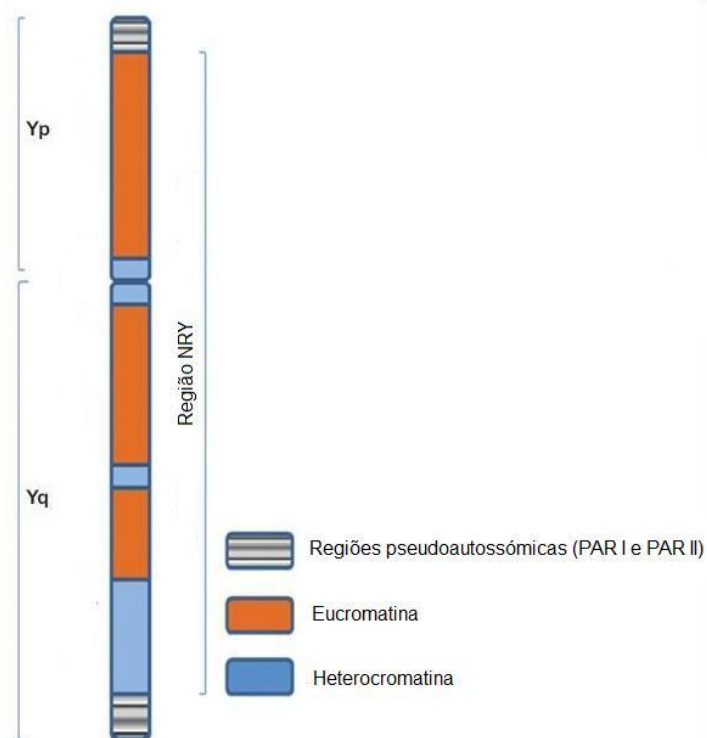


Figura 1 - Estrutura do cromossoma Y humano (Adaptado de Shamsi, Kumar e Dada, 2011)

Para uma plena utilização do potencial do cromossoma Y em termos forenses é fundamental perceber exatamente quais as características que o tornam único. O facto de apenas se encontrar presente em indivíduos do sexo masculino possibilita o seu uso como marcador para identificação do sexo (masculino), através do gene da amelogénina. Da sua amplificação resultam dois fragmentos com 106 pb e 112 pb correspondentes ao cromossoma X e ao cromossoma Y, respetivamente, o que permite a determinação do género de uma amostra biológica (Kashyap *et al.*, 2006). É ainda de destacar que a transmissão da região não recombinante em conjunto de pai para filho aliada às moderadas taxas de mutação da maioria dos polimorfismos desta zona fazem com que homens relacionados pela via paterna, geralmente,

partilhem os mesmos polimorfismos, ou seja, partilhem a mesma linhagem paterna (Kayser e Ballantyne, 2014).

1.3. Marcadores Genéticos

1.3.1. Estrutura e Nomenclatura dos marcadores STRs

Os microssatélites ou *short tandem repeats* (STRs) representam cerca de 4% do genoma humano, sendo elementos mais polimórficos que os *single nucleotide polymorphisms* (SNPs) (Planz e Hall, 2014). Caracterizam-se por serem polimorfismos de comprimento, variando o número e tamanho da unidade de repetição destes marcadores (Figura 2). Uma unidade de repetição é composta por 2-7 pb sendo o número de unidades de repetição mais usados os tri, tetra e pentanucleotídicos, destacando-se especificamente em genética forense os tetranucleotídicos (Primorac, Schanfield e Marjanović, 2014). As principais razões para tal devem-se ao facto de possuírem alelos com tamanhos próximos, permitindo assim uma análise simultânea, por darem origem a produtos de PCR de pequeno tamanho, o que facilita a análise de ADN de amostras degradadas e, por último, verifica-se a existência de uma percentagem mais reduzida de produtos *stutter*, apresentando o fragmento detetado uma altura em unidades de fluorescência relativa (RFUs) igual ou inferior a cerca 15% do fragmento relativo ao alelo que se pretende determinar, valor que aumenta até cerca de 30% usando STRs di e trinucleotídicos. A redução deste fenómeno facilita bastante a interpretação de

resultados provenientes de amostras de mistura de material biológico (Butler, 2005; Pinheiro, 2010).

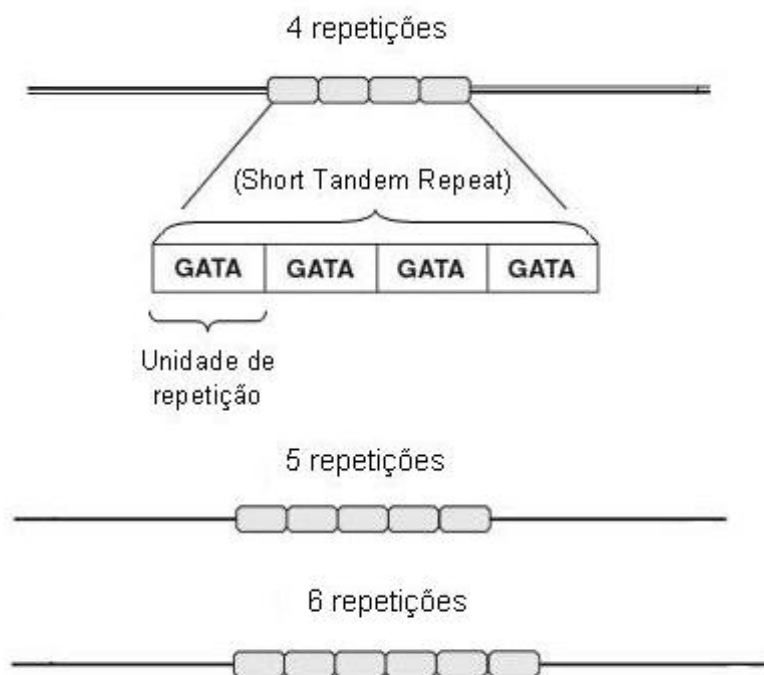


Figura 2 - Estrutura de um STR (Adaptada de Butler, 2010).

Considerando o padrão das unidades de repetição, os STRs podem ser classificados como repetições simples, contendo unidades de repetição idênticas em tamanho e sequência; repetições compostas que incluem duas ou mais repetições simples; repetições complexas que compreendem vários blocos de unidades de repetição cujo tamanho e sequências são variáveis (Urquhart *et al.*, 1994), e por fim, repetições complexas hipervariáveis que contêm inúmeros alelos não consensuais que diferem em tamanho e sequência. Existem alelos, denominados microvariantes, que são constituídos por unidades de repetição incompletas, podendo ter apenas alguns nucleótidos no final da repetição (Butler, 2005).

A Comissão de ADN da Sociedade Internacional de Genética Forense (ISFG) recomendou em 1993 regras de nomenclatura para os marcadores STRs, de forma a facilitar a comunicação e comparação de dados entre laboratórios, desenvolvimento de bases de dados e exercícios de proficiência (Gusmão *et al.*, 2006; Primorac, Schanfield e Marjanović, 2014).

A abundância dos STRs no genoma humano aliada à simplicidade e rapidez de análise laboratorial em sistemas *multiplex* possibilita um elevado grau de individualização. Este conjunto de características fez com que o uso destes marcadores se generalizasse no campo da genética, tornando-se uma ferramenta padrão em estudos populacionais, desenvolvimento de modelos de evolução, estudos de associação com doenças e identificação individual (Planz e Hall, 2014).

1.3.2. Polimorfismos do Cromossoma Y

Os STRs do cromossoma Y (Y-STRs) localizam-se na parte não recombinante do cromossoma (NRY), estando presentes na forma haplóide. Ao longo dos anos os Y-STRs tornaram-se particularmente úteis em estudos evolutivos, genealógicos e forenses (Ballantyne *et al.*, 2010).

O processo de introdução dos Y-STRs na genética forense tem sido longo e demorado, tendo-se iniciado ligeiramente mais tarde do que para os STRs autossómicos. O primeiro Y-STR descrito em 1992 foi a repetição tetranucleotídica do DYS19 (na altura denominada Y27H29), tendo sido imediatamente usado em contexto forense (Roewer e Epplen, 1992). Em meados da década de 90, foi selecionado um conjunto de 9 *loci* (DYS19, DYS389I, DYS389II, DYS390, DYS391,

DYS392, DYS393, DYS285a/b) que constituíam o grupo mínimo de Y-STRs a usar na área forense capazes de diferenciar indivíduos do sexo masculino, de linhagens distintas. Este painel de marcadores ficou conhecido como haplótipo mínimo (Butler, 2006). Em 2003, o SWGDAM (*Scientific Working Group on DNA Analysis Methods*) recomendou o uso de dois Y-STRs adicionais (DYS438 e DYS439), estabelecendo o grupo de base para análise forense nos Estados Unidos, ao qual foram ainda adicionados 6 novos marcadores (DYS437, DYS448, DYS456, DYS458, DYS635, YGATA H4), aumentando assim o poder de discriminação (Ballantyne e Kayser, 2014).

Os Y-STRs que apresentam uma grande diversidade intra e inter populações geralmente partilham determinadas características que conduzem a um aumento das taxas de mutação e, conseqüente, aumento da diversidade alélica. Contribuem para tal o número de repetições por *locus* (quanto maior for o número de repetições maior será a taxa de mutação), a complexidade da sequência do STR (sequências mais complexas em termos de unidades de repetição apresentam taxas de mutação superiores), e o tamanho da unidade de repetição (aumentando a taxa de mutação com a diminuição do tamanho da unidade) (Ballantyne *et al.*, 2010).

Como foi referido anteriormente, a Comissão de ADN da ISFG elaborou recomendações sobre a nomenclatura a utilizar para Y-STRs, para que haja uniformidade na introdução de novos marcadores na área forense, evitando-se assim o uso de diferentes nomenclaturas para um mesmo marcador (Gill *et al.*, 2001; Gusmão *et al.*, 2006).

1.3.3. Taxas de Mutação em Y-STRs

A elevada quantidade de divisões celulares num organismo faz com que a probabilidade de ocorrerem erros seja enorme. A maioria das alterações genéticas que ocorrem resulta de falhas nos mecanismos celulares normais. Erros na duplicação, na recombinação ou na reparação do ADN podem originar alterações permanentes na sequência de ADN, ou seja, mutações (Alberts *et al.*, 2010; Pinto, Gusmão e Amorim, 2014).

As mutações que modificam a sequência nucleotídica podem ser substituições de um par de bases por outro, deleções, inserções, duplicações e translocações, contribuindo todas elas para manter um elevado nível de heterogeneidade na população (Nussbaum, McInnes e Willard, 2007).

As mutações podem ocorrer nas células somáticas e nas células da linha germinativa. As células somáticas constituem os tecidos de um organismo, pelo que alterações nucleotídicas nestas dão origem a células mutadas idênticas, que poderão apresentar um crescimento descontrolado, afetando apenas o próprio indivíduo. Pelo contrário, mutações que ocorram em células da linha germinativa podem ser transmitidas à descendência, tendo estas interesse forense. Algumas mutações apesar de causarem alterações na sequência, não apresentam manifestações funcionais, podendo fixar-se na população, contribuindo assim para alterações evolutivas nos genomas (Alberts *et al.*, 2010).

O que faz com que os STRs sejam marcadores com especial interesse em estudos de genética forense são as suas relativamente elevadas, mas ainda assim suficientemente estáveis, taxas de mutação. Existem vários modelos explicativos

para o aparecimento de eventos mutacionais, sendo o mais aceite a ocorrência de erros durante o processo de replicação da cadeia de ADN, devido ao deslizamento das cadeias durante esta fase, conduzindo à deleção ou inserção de unidades repetitivas (Fan e Chu, 2007).

As taxas de mutação em STRs são influenciadas por diversos fatores, nomeadamente, tamanho, estrutura repetitiva, composição da unidade de repetição e ainda, pela idade e sexo dos progenitores, sendo as taxas de mutação mais elevadas em indivíduos do sexo masculino devido ao maior número de divisões que ocorrem na linha germinal masculina (Fan e Chu, 2007)

Em casos de testes de paternidade, especificamente, há que ter em conta estas taxas para uma correta interpretação dos resultados de Y-STRs, uma vez que uma única diferença alélica entre pai e filho, para o mesmo marcador, não significa necessariamente uma exclusão da paternidade, tendo que ser considerada a possibilidade de ocorrência de uma mutação com base na sua frequência (Pinto, Gusmão e Amorim, 2014).

As taxas de mutação variam não só de marcador para marcador, mas também consoante o grupo populacional em causa. O seu cálculo pode ser feito através de modelos estatísticos, ou a partir da observação direta de casos de investigação de parentesco (Kayser *et al.*, 2000).

Com a colaboração de um elevado número de laboratórios a nível mundial, foi criada uma base de dados *online* disponível no site www.yhrd.org (Y-STR Haplotype Reference Database) que reúne informação acerca não só das taxas de mutação, mas também das frequências haplotípicas, alelos já detetados em populações estudadas, etc.

1.3.4. Kits Comerciais de marcadores genéticos

Presentemente, os STRs são os marcadores de eleição em investigação forense devido à sua considerável variabilidade constatada através de um grande número de *loci*. O *Forensic Science Service* (FSS) do Reino Unido foi das primeiras instituições a desenvolver um *multiplex* para fins forenses, composto por 4 marcadores (Butler, 2006). O esforço conjunto de diversos laboratórios conduziu à seleção de um conjunto de marcadores STRs, denominados por “*core loci*”, que visam facilitar a comparação de resultados entre diferentes laboratórios e a construção de bases de dados nacionais e internacionais, assegurando assim a uniformidade de metodologias. Ao longo dos anos foram desenvolvidos uma grande variedade de *kits* comerciais para dar resposta às diferentes necessidades de um laboratório forense. A utilização destes *kits* possibilita não só a genotipagem de um elevado número de marcadores em simultâneo, obtendo-se assim mais informação numa só reação, como minimiza a ocorrência de erros técnicos pelo facto de ser composto por reagentes otimizados (Mulero e Hennessy, 2014).

Entre os *kits* disponíveis no mercado para o estudo dos Y-STRs encontram-se o *PowerPlex® Y*, o *AmpFISTR® Yfiler®*, o *PowerPlex® Y23* e o *AmpFISTR® Yfiler® Plus*, que possibilitam, respetivamente, a análise de 12, 17, 23 e 27 marcadores. Estão disponíveis comercialmente *multiplexes* suplementares, que visam a tipagem de um conjunto de *loci* adicionais, porém não são frequentemente utilizados na rotina laboratorial (Diegoli, 2015) . Com o aumento considerável de Y-STRs disponíveis, incluindo informação acerca das suas taxas de mutação e diversidade, seria possível seleccionar o conjunto de marcadores adequado à obtenção dos melhores

resultados entre diversas populações, consoante o objetivo inicial (Ballantyne e Kayser, 2014).

1.3.4.1. Sistema AmpFISTR® Yfiler®

O kit AmpFISTR® Yfiler® permite, como referido anteriormente, a amplificação simultânea de 17 Y-STRs, nos quais se incluem os *loci* que constituem o haplótipo mínimo (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393), os recomendados pelo SWGDAM (*Scientific Working Group on DNA Analysis Methods*) (DYS438, DYS439) e ainda seis marcadores altamente polimórficos (DYS437, DYS448, DYS456, DYS458, Y GATA H4, DYS635) (*Applied Biosystems*, 2012). Estes marcadores apresentam produtos de amplificação com tamanhos situados entre os 103 pb e os 327 pb. Os *loci* estão marcados com cinco fluorocromos diferentes (6-FAM™, NED™, PET®, VIC®, e LIZ®), os quais permitiram o aumento do limite de deteção para os 660nm, possibilitando assim a inclusão de *loci* adicionais neste *multiplex* (Mulero *et al.*, 2006).

Na tabela 1 estão indicadas algumas das características dos 17 *loci* presentes no kit AmpFISTR® Yfiler®, nomeadamente, designação do *locus*, número de alelos, estrutura da unidade de repetição e taxa de mutação. O número de repetições apresentado para cada unidade de repetição não corresponde a todos os alelos até agora descritos, mas sim aos presentes no *ladder* deste *kit*.

Tabela 1 - Marcadores incluídos no *kit AmpFISTR® Yfiler®* e respectivos alelos, estrutura repetitiva e taxa de mutação (Obtido no *site* <http://yhrd.org/>, visualizado em Maio de 2015).

Locus	Alelos	Motivo repetitivo	Taxa de mutação (x10⁻³)
DYS456	13-18	(AGAT) _n	4.29
DYS389 I	10-15	(TCTG) ₃ (TCTA) _n	2.93
DYS390	18-27	(CTAT)(CTGT) _n (CTAT) _n (CTGT) ₀₋₁ (CTAT) ₀₋₄ CTA (CTAT) ₂	2.11
DYS389 II	24-34	(TCTG) _n (TCTA) _n n ₂₈ (TCTG) ₃ (TCTA) _n	4.12
DYS458	14-20	(gaaa) ₃ aggaag (GAAA) _n	6.36
DYS19	10-19	(taga) ₃ tagg (TAGA) _n	2.24
DYS385 a/b	7-25	(aagg) ₆₋₇ (GAAA) _n	2.45
DYS393	8-16	(AGAT) _n	1.05
DYS391	7-13	(tactg) ₃ (TCTA) _n	2.45
DYS439	8-15	(GATA) _n	5.45
DYS635	20-26	(TCTA) ₄ (TGTA) ₂ (TCTA) ₂ (TGTA) ₂ (TCTA) ₂ (TGTA) _{0,2} (TCTA) _n	4.33
DYS392	7-18	(TAT) _n	5.19
Y GATA H4	8-13	(AGAT) ₄ CTAT(AGAT) ₂ (AGGT) ₃ (AGAT) _n n ₂₄ (ATAG) ₄ (ATAC) ₁ (ATAG) ₂	3.03
DYS437	13-17	(TCTA) _n (TCTG) ₁₋₃ (TCTA) ₄	1.22
DYS438	8-13	(TTTTTC) ₁ (TTTTTA) ₀₋₁ (TTTTTC) _n	3.75
DYS448	17-24	(AGAGAT) _n agagagatag (AGAGAT) ₃ n ₁₄ (AGAGAT) _n	1.52

1.3.4.2. Sistema PowerPlex® Y23

Em 2012, começou a ser comercializado um novo *kit* para o estudo de 23 Y-STRs, desenvolvido com base nos avanços científicos no campo dos marcadores do cromossoma Y.

O PowerPlex® Y23 combina os 17 *loci* presentes no *multiplex AmpFISTR® Yfiler®* com seis novos Y-STRs altamente discriminativos (DYS481, DYS533, DYS549, DYS570, DYS576 e DYS643) (Thompson *et al.*, 2013). Estes novos *loci* possuem uma diversidade génica superior à maioria dos *loci* presentes noutros *kits* comerciais, possibilitando um maior poder de discriminação entre indivíduos do sexo masculino não relacionados. Uma das razões para este poder acrescido deve-se às taxas de mutação relativamente altas destes seis Y-STRs, que contribuem para o aumento da diversidade alélica, e consequentemente uma melhor capacidade para diferenciar indivíduos da mesma linhagem. Contudo, o nível de taxa de mutação tem que ser tido em conta quando se tratam de casos de paternidade e de identificação de restos cadavéricos, onde se pode recorrer a familiares (Thompson *et al.*, 2012). Além disso, o maior grau de polimorfismo dos marcadores DYS570 e DYS576, que se estima ser superior a 1%, conduz a um aumento do poder de discriminação entre indivíduos de uma mesma população ou relacionados entre si (Thompson *et al.*, 2013).

Na tabela 2 está representada a informação genética relativa aos Y-STRs que constituem este *multiplex*, destacando-se a sombreado os seis novos marcadores incluídos com o objetivo de incrementar o poder de discriminação (Geppert, Edelman e Lessig, 2009).

Tabela 2 - Marcadores incluídos no *kit PowerPlex® Y23* e respectivos alelos, estrutura repetitiva e taxa de mutação (Obtido no *site* <http://yhrd.org/>, visualizado em Maio de 2015).

Locus	Alelos	Motivo repetitivo	Taxa de mutação
DYS456	13-18	(AGAT) _n	4.29 x 10 ⁻³
DYS389 I	10-15	(TCTG) ₃ (TCTA) _n	2.93 x 10 ⁻³
DYS390	18-27	(CTAT)(CTGT) _n (CTAT) _n (CTGT) ₀₋₁ (CTAT) ₀₋₄ CTA (CTAT) ₂	2.11 x 10 ⁻³
DYS389 II	24-34	(TCTG) _n (TCTA) _n n ₂₈ (TCTG) ₃ (TCTA) _n	4.12 x 10 ⁻³
DYS458	14-20	(gaaa) ₃ aggaag (GAAA) _n	6.36 x 10 ⁻³
DYS19	10-19	(taga) ₃ tagg (TAGA) _n	2.24 x 10 ⁻³
DYS385 a/b	7-25	(aagg) ₆₋₇ (GAAA) _n	2.45 x 10 ⁻³
DYS393	8-16	(AGAT) _n	1.05 x 10 ⁻³
DYS391	7-13	(tactg) ₃ (TCTA) _n	2.45 x 10 ⁻³
DYS439	8-15	(GATA) _n	5.45 x 10 ⁻³
DYS635	20-26	(TCTA) ₄ (TGTA) ₂ (TCTA) ₂ (TGTA) ₂ (TCTA) ₂ (TGTA) _{0,2} (TCTA) _n	4.33 x 10 ⁻³
DYS392	7-18	(TAT) _n	5.19 x 10 ⁻³
Y GATA H4	8-13	(AGAT) ₄ CTAT(AGAT) ₂ (AGGT) ₃ (AGAT) _n n ₂₄ (ATAG) ₄ (ATAC) ₁ (ATAG) ₂	3.03 x 10 ⁻³
DYS437	13-17	(TCTA) _n (TCTG) ₁₋₃ (TCTA) ₄	1.22 x 10 ⁻³
DYS438	8-13	(TTTTTC) ₁ (TTTTTA) ₀₋₁ (TTTTTC) _n	3.75 x 10 ⁻³
DYS448	17-24	(AGAGAT) _n agagagatag (AGAGAT) ₃ n ₁₄ (AGAGAT) _n	1.52 x 10 ⁻³
DYS643	6-17	(CTTTT) _n	2.24 x 10 ⁻³
DYS570	10-25	(TTTC) _n	1.79 x 10 ⁻²
DYS481	17-32	(CTT) _n	6.71 x 10 ⁻³
DYS576	11-23	(AAAG) _n	1.79 x 10 ⁻²
DYS549	7-17	(GATA) _n	2.24 x 10 ⁻³
DYS533	7-17	(ATCT) _n	0

1.4. Genética populacional e parâmetros estatísticos de interesse forense

Para a realização de um estudo genético populacional é necessário definir, em primeiro lugar, a população em estudo. Neste caso, a população de referência selecionada foi a população do Sul de Portugal. As amostras dos indivíduos do sexo masculino foram selecionadas aleatoriamente a partir de processos de investigação de parentesco provenientes das comarcas da circunscrição médico-legal da região Sul do país.

A determinação das frequências alélicas em diferentes populações constitui uma das formas de avaliar o grande polimorfismo dos marcadores genéticos a utilizar na rotina forense. O seu cálculo é feito através do quociente entre o número de alelos observados na amostra e o número total de indivíduos em estudo. Assim é possível obter uma caracterização da variação genética de cada um dos marcadores na população em estudo. Para que haja uma maior precisão nos resultados, a amostra considerada deverá ser o mais representativa possível da população e os indivíduos não terem qualquer tipo de relação de parentesco entre si (Butler, 2010).

O facto de não ocorrer recombinação entre os marcadores localizados no cromossoma Y faz com que existam algumas diferenças em relação aos marcadores autossómicos. A estimativa das frequências haplotípicas é feita diretamente a partir da razão entre o número de indivíduos observados na amostra com o haplótipo x e o número total de haplótipos distintos na amostra. Este cálculo é portanto independente do número de marcadores analisados, ou seja, da frequência alélica de cada *locus* (Pinheiro, 2010).

A variabilidade genética nos marcadores do cromossoma Y poderá ser estudada através da diversidade génica (GD), parâmetro que indica, para um determinado *locus*, a probabilidade de dois alelos escolhidos ao acaso na população serem diferentes. Esta pode ser calculada usando a seguinte expressão (Purps *et al.*, 2014):

$$GD = \frac{n(1 - \sum p_i^2)}{(n - 1)}$$

n = número total de amostras

p_i = frequência de cada alelo

É ainda de extrema relevância a estimativa da diversidade haplotípica (HD), parâmetro que indica a probabilidade de encontrar dois perfis genéticos diferentes na população. O seu cálculo é feito com base na fórmula:

$$HD = \frac{n(1 - \sum p_i^2)}{(n - 1)}$$

n = número de haplótipos

p_i = frequência de cada haplótipo

Por último, a capacidade de discriminação permite avaliar qual o potencial para diferenciar indivíduos. A sua determinação é feita dividindo o número total de haplótipos diferentes pelo número total de indivíduos analisados (Coble, Hill e Butler, 2013). Quanto mais próximo de 1 for este valor maior a quantidade de haplótipos diferentes obtidos, e consequentemente maior será a capacidade de diferenciar um indivíduo de outro.

Para a construção de bases de dados populacionais é vantajoso analisar não só um grande número de amostras mas também o maior número de Y-STRs possível, tal faz com que o poder de discriminação haplotípico seja mais elevado e os valores obtidos para os diversos parâmetros estatísticos sejam o mais próximos possível da realidade.

1.5. Aplicação Forense dos Y-STRs

A especificidade do cromossoma Y humano torna-o particularmente útil em estudos e investigações forenses. O uso dos marcadores deste cromossoma para os fins mencionados requer a seleção de *loci* polimórficos, a sequenciação da estrutura dos marcadores, a definição da nomenclatura, determinação das taxas de mutação, o estudo de parâmetros com interesse forense, e o desenvolvimento de métodos de tipagem que permitam um melhor desempenho na análise de amostras degradadas ou com pouco ADN, assim como a construção de bases de dados de haplótipos (Carracedo, 2005).

Os Y-STRs são mais frequentemente utilizados em casos de agressão sexual, onde existe uma mistura de material biológico de origem feminina e masculina. A análise deste tipo de amostras é dificultada pela existência de mais que um contribuinte masculino, como sucede em casos de violação múltipla (Hu *et al.*, 2014). Estando o componente feminino presente em maior proporção, torna-se difícil ou até mesmo impossível obter resultados conclusivos utilizando marcadores genéticos autossómicos. Através destes pode não ser possível a deteção do perfil masculino, devido à desproporção dos componentes, o que não acontece quando se

recorre ao estudo dos STRs do cromossoma Y, os quais amplificam apenas a contribuição masculina (Roewer, 2009). Acrescenta-se ainda o facto de a interpretação de um perfil genético correspondente a uma mistura, recorrendo a marcadores localizados no cromossoma sexual masculino, ser bastante mais fácil devido à presença de apenas um alelo por cada um dos marcadores analisados. Se, por exemplo, a amostra em causa contiver ADN da vítima do sexo feminino, do companheiro e do agressor, um perfil genético para STR autossómicos poderá conter até seis alelos por marcador genético, considerando que todos os contribuintes são heterozigóticos. Ao invés, num perfil genético do cromossoma Y apenas irão constar no máximo dois alelos, relativos a cada um dos indivíduos do sexo masculino, visto a fração feminina contida no vestígio não ser amplificada, devido à especificidade dos *primers* (Pinheiro, 2013).

Estes polimorfismos podem também ser utilizados em casos de identificação de restos cadavéricos e em casos de investigação de paternidade quando apenas estão disponíveis amostras biológicas de familiares aparentados pela linhagem paterna com o suposto pai, uma vez que o perfil genético de cromossoma Y será partilhado entre todos os elementos da mesma linhagem. Contudo, este padrão de herança faz com que nenhum dos indivíduos do sexo masculino possa ser excluído da paternidade apenas com base em marcadores do Y, sendo ainda a sua aplicabilidade limitada nos casos de paternidade, pois só poderá ser empregue em casos de descendência masculina. Estes fatores fazem com que, na maioria dos casos, sejam utilizados preferencialmente marcadores autossómicos para que o grau de discriminação não seja condicionado (Ballantyne e Kayser, 2014; Butler, 2005).

É de salientar uma das limitações destes polimorfismos, quer nos casos civis quer nos casos criminais, que reside no facto de não se limitar a um indivíduo em concreto, mas a toda a sua linhagem paterna.

Encontram-se descritos mais de 400 STRs do cromossoma Y, contudo apenas existe informação detalhada acerca da sua sequência, taxas de mutação e diversidade para cerca de 180 marcadores. Os restantes, por ainda não se encontrarem bem caracterizados não são aplicados, por enquanto, à investigação genética forense (Hanson e Ballantyne, 2006).

1.6. Validação de métodos

Num laboratório de Genética Forense existem inúmeros requisitos, de gestão e técnicos, que devem ser implementados de forma a garantir a qualidade do mesmo. Entre eles encontra-se a validação de métodos, um processo através do qual se demonstra que o desempenho do laboratório é robusto, seguro e reproduzível. Está definida na norma ISO/IEC 17025 a lista de métodos que devem ser validados, nomeadamente, a validação de métodos não normalizados, de métodos criados ou desenvolvidos pelo próprio laboratório, de métodos normalizados utilizados fora do seu âmbito de utilização e extensões ou modificações de métodos normalizados. A validação dos métodos visa a confirmação da sua adequabilidade.

A forma de validação irá variar consoante o procedimento de ensaio em análise, dependendo das características do mesmo, complexidade da amostra e

experiência prévia do laboratório na utilização do método, sendo preferível a seleção de métodos de ensaio cujo desempenho seja já conhecido pelo laboratório.

O processo de validação deverá ser executado de duas formas diferentes, através de uma avaliação indireta e/ou uma avaliação direta, realizando os estudos a seguir referidos. A avaliação indireta visa demonstrar as características de desempenho relevantes para o ensaio considerando o âmbito de aplicação previsto, sendo para tal analisados seis parâmetros a seguir descritos.

Em técnicas desenvolvidas para genotipagem de ADN humano deve ser avaliado o seu potencial para detetar ADN de diferentes espécies animais, ou seja, a sua especificidade. Deverá ser possível caracterizar e identificar, de forma inequívoca, ADN humano, sendo negativa a análise de todas as outras amostras de espécies animais.

Quando adequado, deverá ser determinada a quantidade mínima e máxima de ADN numa amostra que permite a obtenção de resultados seguros, aferindo-se assim os limiares analíticos. Para tal, é feita quantificação das amostras seguida de diluições seriadas e nova quantificação da quantidade de ADN presente nas sucessivas diluições. Quando a quantidade de ADN adicionada não é adequada, poderão surgir efeitos estocásticos, como fragmentos adicionais (*stutter*) e adenilação incompleta, caso haja excesso de ADN. Por outro lado, em amostras com pouco ADN poderá verificar-se um desequilíbrio entre o tamanho de dois alelos num indivíduo heterozigótico e até a perda de um alelo (*allele dropout*). O estudo deste parâmetro permite a longo prazo alcançar uma maior eficiência, pois, em

princípio, irão ser obtidos perfis equilibrados, facilitando assim a sua interpretação e diminuindo a necessidade de repetições.

Entende-se por precisão do método a capacidade de se obter resultados idênticos para as mesmas amostras, usando o mesmo método. Neste estudo é tido em conta o tamanho dos fragmentos obtidos, devendo todos os alelos encontrarem-se localizados a $\pm 0.5\text{pb}$ em relação ao alelo correspondente no *ladder* alélico. Quando tal não ocorre, as amostras deverão ser novamente analisadas para que possa ser feita a distinção entre um alelo *off-ladder* e um erro na atribuição alélica. Tem portanto que existir concordância entre os resultados realizados entre ensaios independentes. Este parâmetro é avaliado através da repetibilidade e da reprodutibilidade. Um método apresenta boa repetibilidade quando por análise em duplicado de uma mesma amostra, em condições idênticas, se obtêm dois perfis genéticos iguais. Para garantir a coerência dos resultados avalia-se ainda a reprodutibilidade, efetuando-se o estudo das amostras intralaboratorialmente, por diferentes técnicos, ou interlaboratorialmente, recorrendo ao mesmo procedimento.

Por último, o laboratório deve demonstrar que o procedimento consegue discriminar amostras contaminadas de amostras não contaminadas. O facto de os novos *kits* comerciais serem cada vez mais sensíveis faz com que a probabilidade de contaminação seja maior, visto poderem ser amplificadas quantidades exíguas de ADN exógeno. As contaminações podem ocorrer devido a erros de pipetagem em diversas fases do processo, e durante a colheita e manuseamento das amostras, podendo resultar em erros na análise dos resultados. As contaminações podem pôr em causa a integridade dos resultados, devendo por isso ser feita a análise de controlos negativos.

No que diz respeito à avaliação direta, esta visa uma comparação do procedimento a validar com métodos normalizados ou de referência, recorrendo a padrões e materiais de referência ou através da obtenção e reconhecimento da eficiência do método a nível internacional. A comparação com métodos desenvolvidos é um dos parâmetros considerados, onde o procedimento a executar apresenta ligeiras alterações em relação ao protocolo descrito no manual de utilização do *kit* comercial em questão. Os resultados obtidos com ambas as técnicas devem ser semelhantes, apesar das modificações introduzidas.

As comparações interlaboratoriais são essenciais para demonstrar a coerência dos resultados, devendo o laboratório participar em ensaios de avaliações externas da qualidade.

É ainda extremamente importante a utilização de materiais de referência como padrão de comparação dos resultados. Estes materiais caracterizam-se pela sua homogeneidade e estabilidade, estando aptos para o uso pretendido num processo de medição. Todos os instrumentos e equipamentos deverão ser sujeitos a testes de controlo e calibração antes da sua utilização na rotina.

Todos os ensaios acima referidos devem ser realizados pelo laboratório antes da introdução de uma nova metodologia, garantindo assim a sua competência na genotipagem de ADN. Todos os procedimentos técnicos devem ser cuidadosamente documentados assegurando que cada amostra é corretamente manuseada e processada.

1.7. Objetivos do estudo

Neste trabalho foi efetuado o estudo das frequências alélicas e haplotípicas dos 23 marcadores do cromossoma Y presentes no *kit PowerPlex® Y23* e calculados os respetivos parâmetros estatísticos forenses, em 243 amostras da população da região do Sul de Portugal, provenientes de investigações de paternidade efetuadas no Serviço de Genética e Biologia Forenses do INMLCF, I.P. – Delegação do Sul (SGBF-S), com o objetivo da sua aplicação na atividade pericial desenvolvida neste Serviço.

Paralelamente ao referido, foi desenvolvido um estudo de validação do *kit PowerPlex® Y23*, no qual foi efetuada uma análise comparativa utilizando 66 amostras da região Sul de Portugal, já tipadas com o *kit PowerPlex® Y23* e realizada a sua análise com o *kit AmpFISTR® YFiler®*, no sentido de verificar qual a diferença ao nível do poder de discriminação dos dois kits. Foi ainda avaliada a influência da variação do volume final da reação de PCR na obtenção das tipagens e determinados os valores de *stutter* e desequilíbrio alélico no marcador DYS385a/b nas 243 amostras em estudo.

II - Materiais e Métodos

2.1. Amostragem

Neste estudo foram analisadas 243 amostras de indivíduos do sexo masculino, saudáveis e não aparentados na região Sul de Portugal. Estas amostras são provenientes de investigações de paternidade efetuadas no SGBF-S, o qual abrange os distritos de Beja, Évora, Faro, Lisboa, Portalegre, Santarém e Setúbal.

As amostras, anonimizadas, foram obtidas no âmbito das perícias realizadas no INMLCF, ao abrigo do artigo 31º da Deliberação nº 849/2010, de 7 de maio (Regulamento Interno do INMLCF), que prevê a utilização de amostras de rotina pericial para fins de investigação científica, pelo que não suscita questões de natureza ética ou legal.

2.2. Procedimento Experimental

O procedimento experimental realizado compreende diversas etapas laboratoriais, as quais se encontram esquematicamente representadas no seguinte fluxograma.

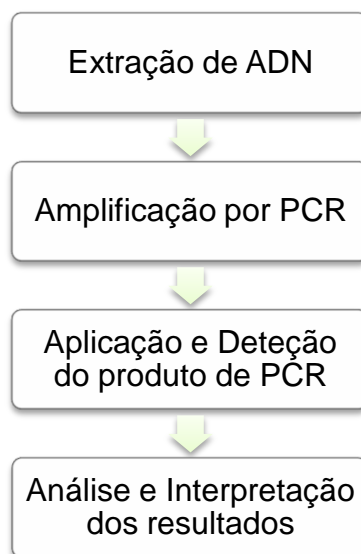


Figura 3 - Fluxograma representativo das fases do procedimento experimental.

2.2.1. Extração de ADN

A extração de ADN a partir de manchas de sangue foi realizada pelo método de Chelex[®] 100. O Chelex é uma resina quelante com grande afinidade para iões metálicos, como os de magnésio que se encontram presentes nas células. O uso desta resina evita a degradação do ADN, devido ao seu efeito quelante sobre os iões metálicos que poderão atuar como catalisadores na desnaturação do ADN a altas temperaturas (Walsh, Metzger e Higuchi, 2013).

Para a preparação da solução de Chelex a 5%, diluiu-se 2.5g de Chelex em 50ml de água desionizada.

Para cada amostra foi retirado um círculo da mancha de sangue armazenada em cartão *WhatmanTM FTATM*, com um *paper punch single hole*, colocando-se o mesmo num *eppendorf* de 1.5mL previamente identificado, ao qual se adicionou 1mL de água desionizada. Incubaram-se os tubos à temperatura ambiente, durante

aproximadamente 15 minutos, seguindo-se uma agitação no *vortex* a baixa rotação durante 5-10 segundos. Centrifugou-se os tubos a 14000 rotações por minuto (rpm) durante 3 minutos a uma temperatura de 20°C e retirou-se com micropipeta cerca de 980µL de sobrenadante de cada amostra.

Ao volume restante de *pellet* acrescentou-se 180µL de solução de Chelex a 5% previamente preparada e colocaram-se os tubos no banho de aquecimento a 56°C, durante 15 minutos. Seguiu-se nova agitação no *vortex* a alta velocidade, durante 10 a 15 segundos, após a qual se colocaram os tubos em suportes flutuantes num banho em ebulição durante 8 minutos. Efetuou-se uma última agitação no *vortex* a alta velocidade durante alguns segundos e centrifugou-se novamente a 14000rpm durante 5 minutos. Por fim, as amostras foram armazenadas no congelador a -20°C.

Em cada extração foi preparado um controlo negativo de acordo com o procedimento acima descrito, mas contendo água desionizada em vez de amostra.

2.2.2. Amplificação por PCR

A amplificação de ADN é realizada pela técnica de *Polimerase Chain Reaction* (PCR), através da qual se obtêm um elevado número de cópias de um fragmento de ADN de interesse. Para tal, dois *primers* irão delimitar a região a amplificar, hibridando com as extremidades 3' de cada uma das cadeias polinucleotídicas. Este processo é composto por três fases que se repetem por um determinado número de ciclos que varia entre 25-35. Em cada uma das fases ocorrem variações de temperatura, na primeira a cerca de 94°C dá-se a separação

da cadeia dupla – fase de desnaturação. Segue-se uma descida da temperatura até cerca de 60°C que permite a ligação dos *primers* às extremidades 3´da dupla cadeia de ADN – fase de *annealing*. Por fim, a temperatura sobe novamente até cerca de 72°C onde por ação da enzima ADN polimerase são sintetizadas novas cadeias – fase de extensão dos *primers* pela adição de didesoxinucleótidos trifosfatados (ddNTPs) (Pinheiro, 2008, 2010).

Para o estudo de STRs do cromossoma Y foi utilizado o *kit* comercial *PowerPlex® Y23* (Promega), o qual permite a coamplificação de 23 marcadores. Todo o procedimento foi efetuado dentro de uma câmara de fluxo laminar, preparando-se em primeiro lugar uma mistura de reação composta por *Master Mix* (que contém ADN polimerase, dNTPs, MgCl₂), *Primer Pair Mix* e água desionizada cujas quantidades por amostra se encontram discriminadas na Tabela 3. A mistura foi realizada para um total de n+1 amostras (onde “n” representa o número de amostras a amplificar), distribuindo-se 11.5µL desta por cada um dos microtubos de 0.2mL, à qual se adicionou 1 µL de solução de ADN a amplificar, perfazendo um volume final para cada reação de 12.5µL. Preparou-se ainda um controlo positivo e negativo, onde se adicionou 1µL de *2800M Control DNA*, 0.5ng/µL e 1µL de água, respetivamente, em vez da solução de ADN.

Tabela 3 - Reagentes e volumes utilizados por amostra na preparação da mistura de reação para amplificação com o *kit PowerPlex® Y23*.

Reagentes	Volume por reação
<i>Master Mix</i>	2.5µL
<i>Primer Pair Mix</i>	1.25µL
Água	7.75µL
Volume Total	11.5µL

A reação de amplificação foi executada no termociclador *GeneAmp 9700* (Applied Biosystems), de acordo com as condições expressas na Tabela 4. Após esta ter terminado, os tubos das reações de amplificação podem ser armazenados a 4°C ou prosseguir para a fase de análise.

Tabela 4 - Condições da reação de amplificação com o *kit PowerPlex® Y23* num termociclador *GeneAmp 9700*.

Passo Inicial	30 Ciclos			Extensão Final	Último Passo
	Desnaturação	<i>Annealing</i>	Extensão		
96°C 2min	94°C 10 seg	61°C 1min	72°C 30seg	60°C 20min	4°C ∞

2.2.3. Aplicação e Detecção dos produtos de PCR

A análise dos produtos amplificados é feita por eletroforese capilar num sistema *ABI Prism 3130 XL* aplicando um polímero desnaturante (POP-4) num capilar de 36cm. A migração dos vários fragmentos é feita consoante o seu peso molecular, sendo que os de menor tamanho migram mais rapidamente que os fragmentos de maior tamanho. A presença de ddNTPs marcados com fluorocromos permite a deteção dos fragmentos, ao emitirem fluorescência quando excitados por um feixe de *laser* incidente no capilar. Esta fluorescência é representada no eletroforegrama por um pico, cuja cor irá corresponder ao ddNTP excitado.

Deu-se início ao procedimento preparando num microtubo de centrífuga de 1.5mL, uma mistura de eletroforese para Y+2 amostras (onde Y representa o número de amostras a analisar), composta por um *standard* interno (*CC5 ILS 500 Y23*), que permite a atribuição do tamanho em número de pares de base do fragmento de ADN amplificado, e por formamida Hi-Di (agente desnaturante). As quantidades por amostra de cada reagente acima referido encontram-se mencionadas na Tabela 5.

Tabela 5 - Reagentes e volumes utilizados por amostra na preparação da mistura de eletroforese capilar com o *kit PowerPlex® Y23*.

Reagentes	Volume por reação
Formamida Hi-Di	10µL
<i>Size standard (CC5 ILS 500 Y23)</i>	0.5µL
Volume Total	10.5µL

Distribuiu-se 10.5µL da mistura de eletroforese por cada poço da placa *MicroAmp® Optical* (Applied Biosystems), adicionou-se 1µL da respetiva reação de amplificação e num dos poços em vez de produto amplificado adicionou-se 1µL de *ladder* alélico fornecido com o *kit PowerPlex® Y23*, para fins de atribuição alélica a cada fragmento. As amostras foram depois desnaturadas num termociclador *GeneAmp® PCR System 9700* durante 3 minutos a 95°C seguida de um arrefecimento até aos 4°C, por um período de 30 segundos a 3 minutos.

Por último, foi realizada a eletroforese capilar num sequenciador *ABI Prism 3130*, e os dados recolhidos usando o *software Data Collection v4.0* (Applied Biosystems).

2.2.4. Análise de resultados

A análise das amostras foi feita com o *software GeneMapperID-X* (versão 1.2). Foi definido um limiar de detecção dos fragmentos de 50 RFUs, abaixo do qual os fragmentos não são considerados para análise.

Para que um perfil genético das amostras seja determinado com sucesso é essencial que ocorra uma correta separação dos fluorocromos pela matriz aplicada no sequenciador *ABI Prism 3130*. O *ladder* alélico e o *size standard* são igualmente importantes, migrando conjuntamente com as amostras durante a eletroforese, permitindo, respetivamente, a determinação do genótipo final de cada amostra e o dimensionamento dos fragmentos de ADN.

2.3. Validação de Procedimento de Ensaio

2.3.1. Estudo de concordância

A análise dos parâmetros a seguir descritos foi executada de acordo com as indicações presentes no Procedimento Geral de Validação de Procedimentos de Ensaio instituído pelo Serviço de Genética e Biologia Forense (SGBF).

2.3.1.1. Comparação com *AmpFISTR*[®] *Yfiler*[®]

Foram tipadas com o *kit AmpFISTR*[®] *YFiler*[®] 66 amostras já analisadas com o *kit PowerPlex*[®] Y23. Para tal, preparou-se uma mistura de reação constituída por *Master Mix*, *Primer Pair Mix*, ADN polimerase e água desionizada nas quantidades descritas na Tabela 6. Distribuiu-se 11.25µL desta por cada um dos microtubos, aos quais se juntou 1.25µL de solução de ADN a amplificar. Foi realizado um controlo positivo e negativo para cada reação de amplificação, tendo esta decorrido no termociclador *GeneAmp 9700* (Applied Biosystems), segundo as condições da Tabela 7.

Por último, procedeu-se à eletroforese capilar do produto amplificado de acordo com o protocolo acima descrito para o *kit PowerPlex*[®] Y23.

Tabela 6 - Reagentes e volumes utilizados por amostra na preparação da mistura de reação para amplificação com o *kit AmpFISTR*[®] *YFiler*[®].

Reagentes	Volume por reação
<i>Master Mix</i>	4.6µL
<i>Primer Pair Mix</i>	2.5µL
ADN polimerase	0.4µL
Água	3.75µL
Volume Total	11.25µL

Tabela 7 - Condições da reação de amplificação com o *kit PowerPlex® Y23* num termociclador *GeneAmp 9700*.

Passo Inicial	30 Ciclos			Extensão Final	Último Passo
	Desnaturação	Annealing	Extensão		
95°C 11min	94°C 1min	61°C 1min	72°C 1min	60°C 80min	4°C ∞

2.3.1.2. Volume de reação

De acordo com o protocolo do *kit PowerPlex® Y23* é recomendado que o volume final da reação de PCR seja 25µL. Assim, foram analisadas as mesmas 36 amostras sob diferentes condições, conforme as orientações do fabricante e reduzindo o volume final para 12.5µL e para 6.25µL. Desta forma é possível determinar as condições alternativas de amplificação segundo as quais se obtém um perfil genético completo e verificar se a redução de volume não interfere na obtenção das tipagens relativamente ao indicado no manual do *kit*.

2.3.2. Artefactos da PCR

No decorrer da PCR é comum surgirem fragmentos adicionais, geralmente com cerca de 15% da dimensão dos fragmentos relativos aos alelos a determinar.

Estes fragmentos são produtos *stutter*, artefactos alélicos com 4 bases a menos que o alelo correspondente (considerando STRs tetranucleotídicos) que resultam de cópias incompletas de alelos, surgindo um fragmento adicional com uma unidade de repetição a menos ($n-1$), ou a mais ($n+1$), que o alelo a determinar (Gill, Sparkes e Buckleton, 1998).

É durante o processo de amplificação por PCR que se formam as *stutter*, por deslizamento da enzima polimerase. O mecanismo proposto aponta para a formação de uma ansa e deslize de cadeias (de ADN-molde ou *primer*), dando origem à inserção ou deleção de uma unidade de repetição na nova cadeia. Em termos de análise forense, os produtos *stutter* podem comprometer a interpretação de perfis de ADN onde exista uma mistura de material biológico, sendo por isso importante determinar qual a tendência para a sua formação em cada marcador (Brookes *et al.*, 2012).

Para o estudo das *stutters* foram considerados os fragmentos com menos uma, e mais uma unidade de repetição que o alelo a determinar. O seu cálculo foi feito dividindo o tamanho do fragmento *stutter*, em RFUs (*Relative Fluorescence Unit*) pelo tamanho do alelo a determinar.

2.3.3. Balanceamento dos heterozigóticos (DYS385a/b)

No marcador *multilocus* DYS385a/b, o tamanho dos alelos na forma heterozigótica deveria ser idêntico, contudo devido a efeitos estocásticos derivados da reação de PCR, como a amplificação preferencial de um dos alelos, poderá

ocorrer um desequilíbrio entre o tamanho de dois alelos de um mesmo STR. O equilíbrio entre alelos é determinado através da razão entre a altura em RFUs dos fragmentos de um mesmo *locus*, devendo o valor obtido ser superior a 70%, para se considerar que os alelos se encontram bem equilibrados.

III – Resultados e Discussão

3.1. Haplótipos do cromossoma Y

No total das 243 amostras de indivíduos do sexo masculino saudáveis e não aparentados residentes na região Sul de Portugal, analisadas com o *kit PowerPlex® Y23* foram observados 237 haplótipos diferentes, dos quais 231 são únicos e 6 haplótipos são partilhados por dois indivíduos. No anexo 1 encontram-se todos os haplótipos obtidos, destacando-se a sombreado os seis haplótipos que ocorreram em dois indivíduos. Posteriormente foram determinadas as respetivas frequências alélicas entre outros parâmetros estatísticos.

3.2. Análise Estatística

Para cada um dos 23 Y-STRs analisados foi efetuado o estudo da variabilidade alélica detetada em cada marcador. Em média, por cada marcador com uma única cópia alélica, ou seja, com exceção do marcador DYS385a/b, verifica-se a ocorrência de 7 alelos. Na figura 4 encontra-se representado o número de características alélicas detetadas em cada Y-STR. A sua análise mostra que nos marcadores com um único *locus*, o número mínimo de alelos observados é 4, nos *loci* DYS389 I, DYS391 e YGATAH4, sendo o número máximo 12 no *locus* DYS458. Estes valores permitem inferir acerca do grau de polimorfismo de cada marcador sendo o DYS458 o mais polimórfico. Quanto ao marcador *multilocus* DYS385a/b foram obtidos 41 haplótipos diferentes. Da sua amplificação resultam dois

fragmentos designados por DYS385a e DYS385b, contudo quando esta é realizada por um único par de *primers* não é possível fazer a distinção entre as duas regiões. Por esta razão os resultados obtidos neste STR são designados por genótipos (Gusmão *et al.*, 2006).

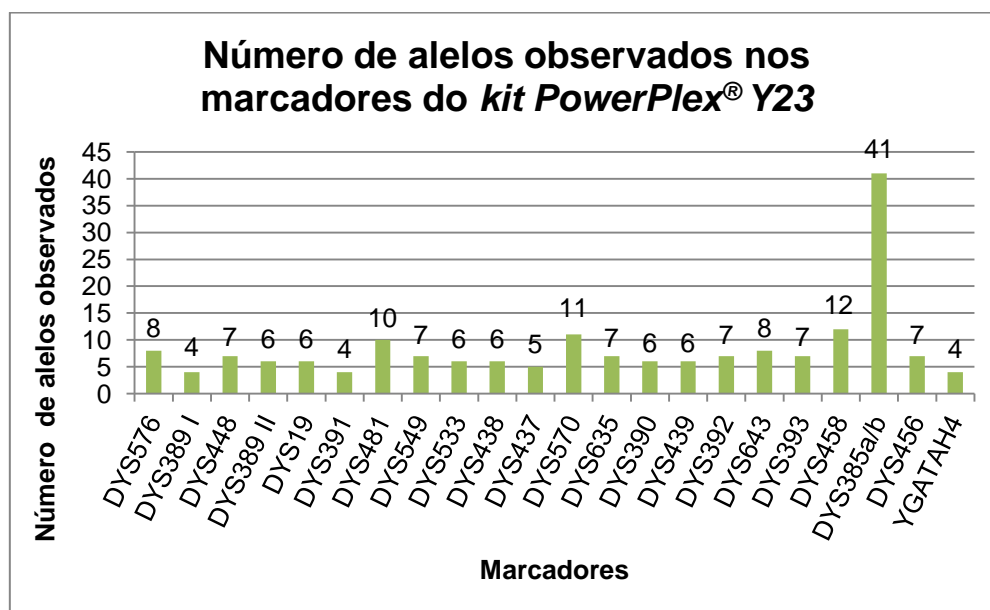


Figura 4 - Número de alelos detetados em cada marcador genético estudado pelo *kit PowerPlex® Y23*

As frequências alélicas de cada *locus* e as frequências genotípicas do marcador DYS385a/b encontram-se representadas na figura 5 e 6, respetivamente, e discriminadas nos anexos 2 e 3, respetivamente. A sua observação conjunta revela uma distribuição de acordo com a curva de Gauss das frequências alélicas em cada marcador, estando um dos alelos sobre representado em comparação com os restantes. São exceção os marcadores DYS385a/b e DYS392, apresentando este último dois alelos, 11 e 13, com frequências bastante elevadas, 0.4568 e 0.4609,

respetivamente. Quanto ao marcador DYS385a/b, dos 41 genótipos distintos observados destaca-se o genótipo 11-14 com uma frequência de 0.2881, representando aproximadamente 30% dos indivíduos do sexo masculino na amostra em estudo. Do total dos genótipos, 12 ocorreram em apenas um indivíduo e os restantes 29 são comuns a mais do que um indivíduo.

No marcador DYS448 foi observada uma duplicação alélica, ou seja, determinada região do STR encontra-se duplicada no próprio *locus*. A duplicação 18-19 observada apresenta uma frequência de 0.0041. Este resultado foi comparado com a base de dados de haplótipos do cromossoma Y (www.yhrd.org), onde a frequência apresenta um valor de 0.0022, o qual é inferior ao obtido na população em estudo, tendo, muito provavelmente, influência no mesmo o tamanho da amostra em análise.

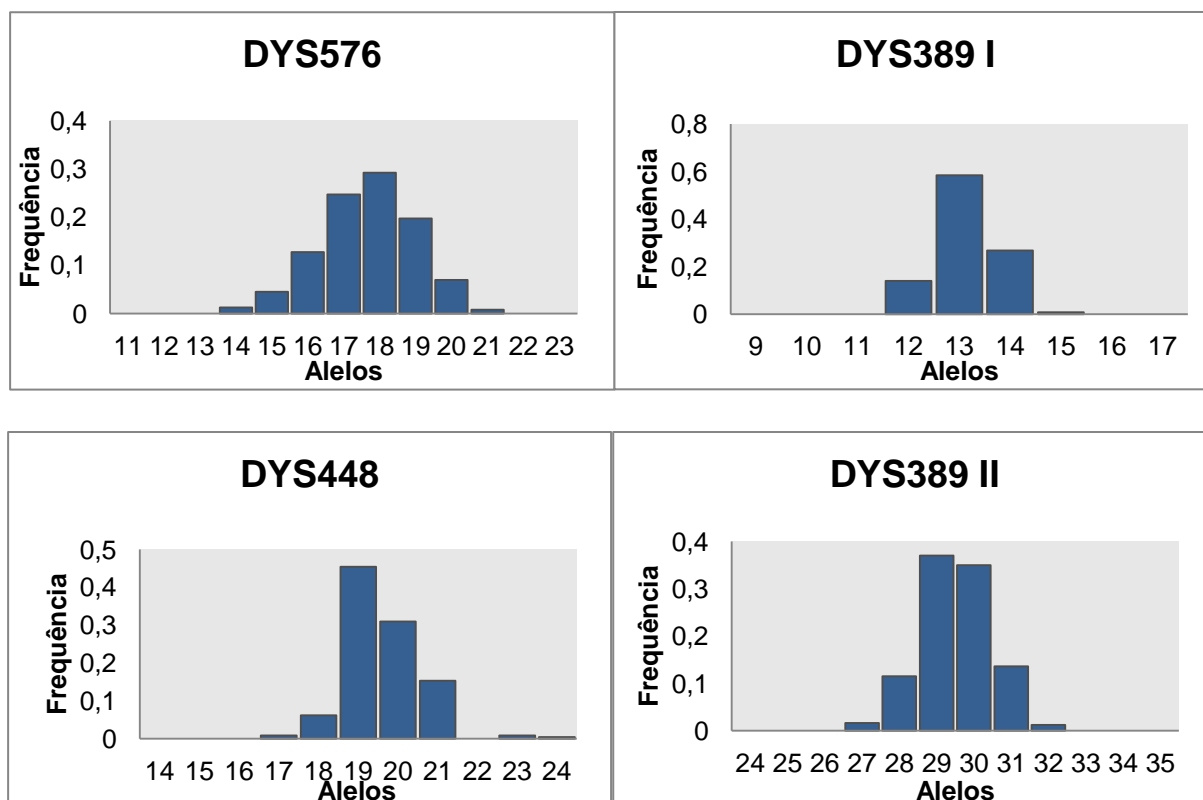


Figura 5 - Distribuição alélica de cada marcador analisado com o *kit PowerPlex® Y23*, nos 243 indivíduos estudados.

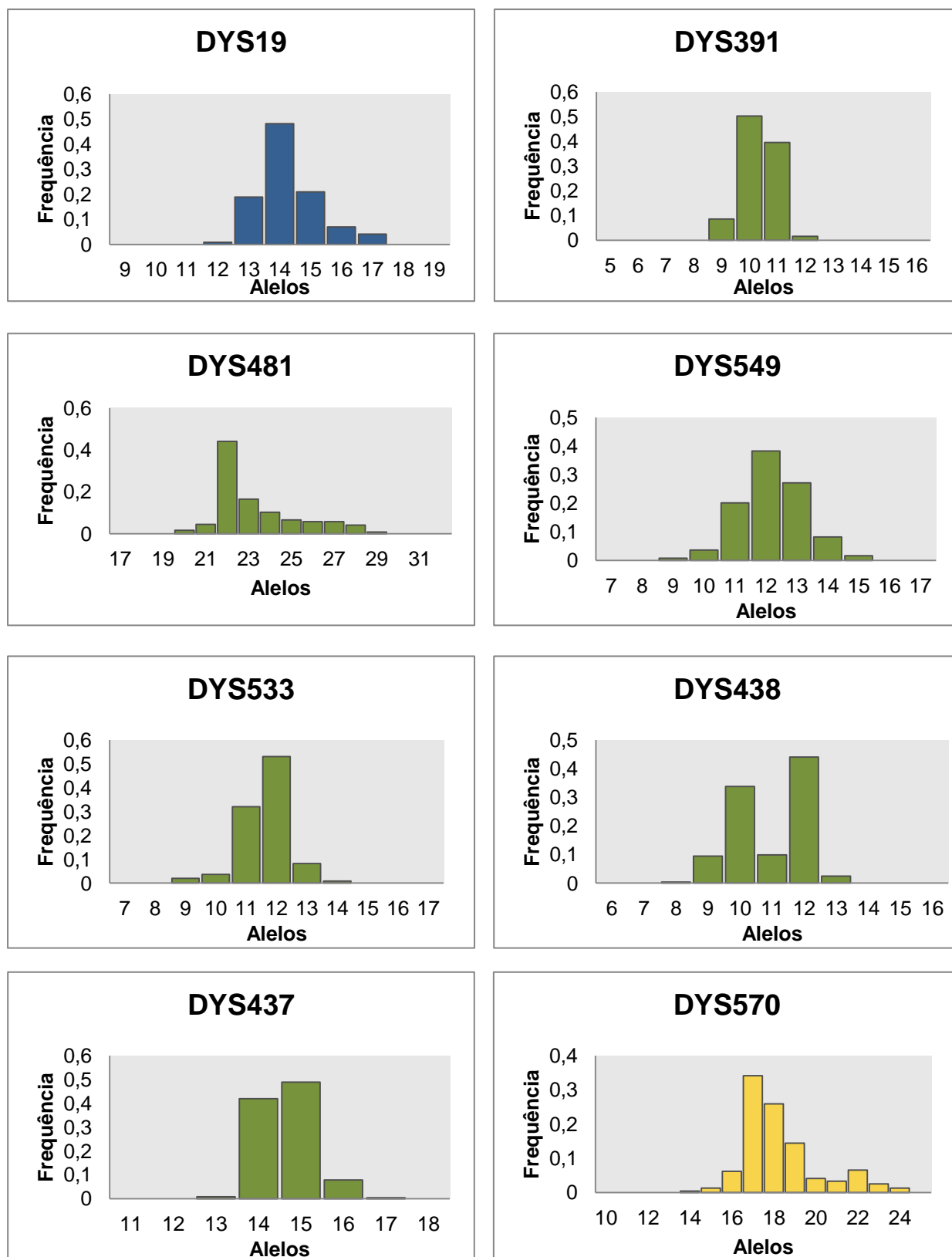


Figura 5 (continuação) - Distribuição alélica de cada marcador analisado com o *kit* PowerPlex® Y23, nos 243 indivíduos estudados.

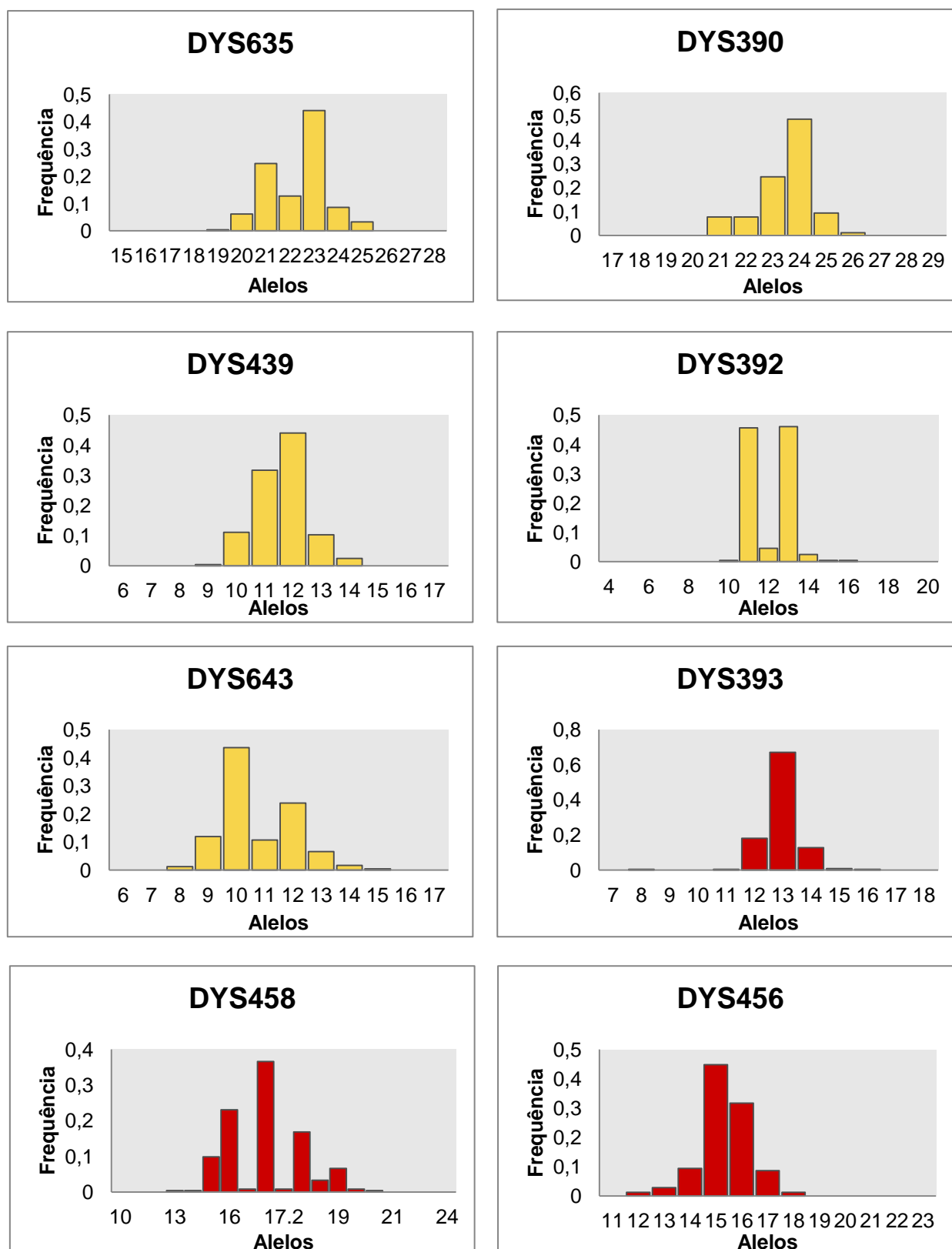


Figura 5 (continuação) - Distribuição alélica de cada marcador analisado com o *kit* PowerPlex® Y23, nos 243 indivíduos estudados.

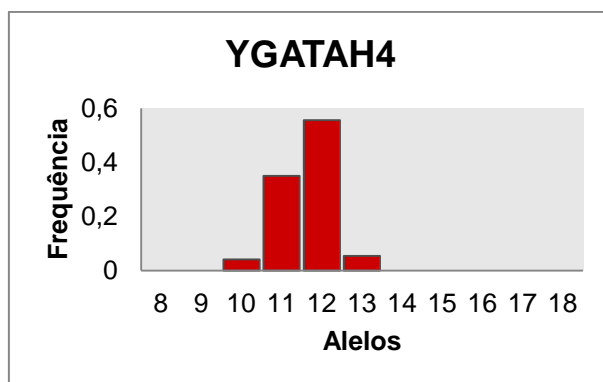


Figura 5 (continuação) - Distribuição alélica de cada marcador analisado com o *kit* PowerPlex® Y23, nos 243 indivíduos estudados.

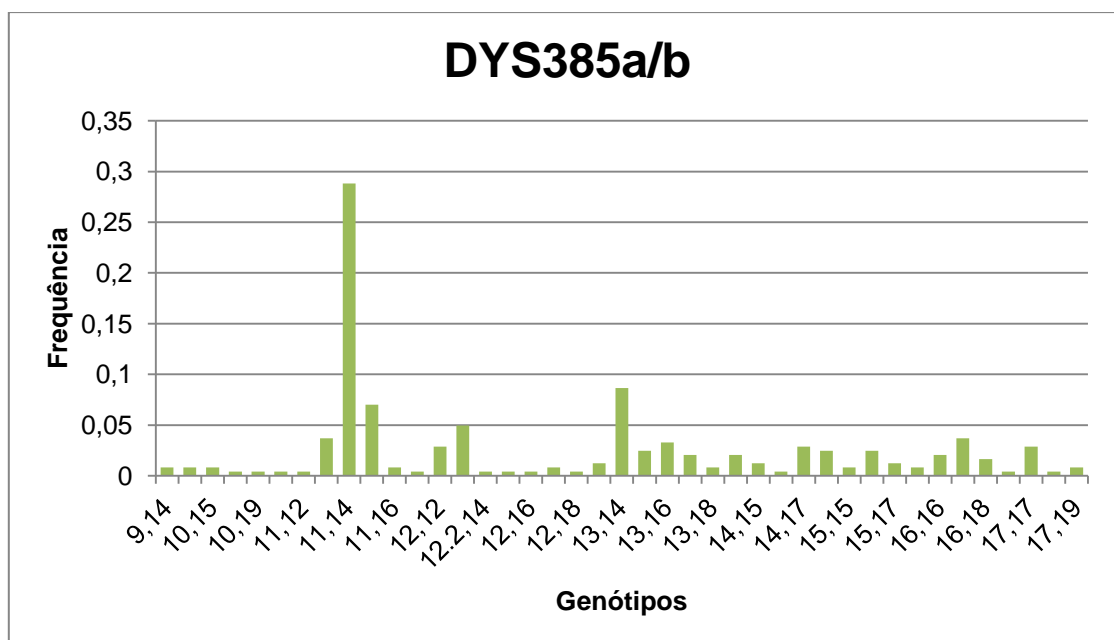


Figura 6 - Frequência genotípica do marcador DYS385a/b.

A análise dos resultados do cálculo da diversidade génica para cada um dos *loci* estudados (tabela 9) mostra uma variação entre 0.5009, para o marcador DYS393, e 0.7912, para o marcador DYS576. No caso do marcador DYS385a/b, este parâmetro não foi calculado por se tratar de um *multilocus*. No conjunto dos 23 Y-STRs analisados, destacam-se os loci DYS576 e DYS570, com os valores de diversidade génica mais elevados. Estes fazem parte dos seis novos Y-STRs altamente discriminativos incluídos no *kit PowerPlex® Y23*, apresentando igualmente as taxas de mutação mais elevadas. Por oposição, entre os seis marcadores suplementares, o *locus* DYS533 revelou menor diversidade génica. Neste marcador foram detetados 6 alelos diferentes, tendo o alelo 12 uma frequência alélica relativamente elevada (anexo 2), sendo por isso pouco discriminativo dos indivíduos do sexo masculino. Ao considerar todos os Y-STRs presentes no *kit*, o menor valor foi observado no *locus* DYS393.

Tabela 8 - Diversidade génica.

Locus	Diversidade Génica
DYS576	0.7912
DYS389I	0.5673
DYS448	0.6700
DYS389II	0.7083
DYS19	0.6816
DYS391	0.5841
DYS481	0.7534
DYS549	0.7306
DYS533	0.6065
DYS438	0.6693
DYS437	0.5778
DYS570	0.7835
DYS635	0.7165
DYS390	0.6779
DYS439	0.6821
DYS392	0.5762
DYS643	0.7223
DYS393	0.5009
DYS458	0.7689
DYS456	0.6808
YGATAH4	0.5644

No que diz respeito à diversidade haplotípica esta foi determinada para o conjunto dos 23 Y-STRs, tendo sido realizada a análise do número total de haplótipos obtidos (anexo 1), nomeadamente, 237 haplótipos diferentes. Para este parâmetro estatístico foi calculado um valor bastante elevado, designadamente 0.999796, tal deve-se ao facto de apenas seis haplótipos serem partilhados por dois

indivíduos. Assim, a probabilidade de dois indivíduos não aparentados na população Sul de Portugal, partilharem o mesmo haplótipo, é de apenas 0.000204.

A diversidade haplotípica desta população ($HD=0.999796$) foi comparada com a de uma população do nordeste de Itália ($HD=0.999994$) (Turrina *et al.*, 2014) e de um conjunto de populações europeias ($HD=0.999992$) (Purps *et al.*, 2014), e apesar de menor enquadra-se nos valores obtidos.

Do estudo de 243 amostras, foram observados 237 haplótipos diferentes, dos quais 231 são únicos e 6 haplótipos são partilhados por dois indivíduos. Assim, no total foram obtidos 234 haplótipos diferentes, tendo a capacidade de discriminação calculada um valor de 0.9630. O aumento do número de Y-STRs analisados, e de marcadores com maior diversidade génica tem influência neste parâmetro, sendo expectável um aumento da capacidade de discriminação.

Tem que ser considerado, para todos os parâmetros estatísticos acima referidos, o facto de existirem fluxos populacionais, designadamente fenómenos de imigração, que levam à introdução de novas características genéticas na população e, conseqüentemente, à alteração do equilíbrio populacional.

3.3. Variantes alélicas

Nas 243 amostras analisadas foram identificadas um total de cinco microvariantes alélicas, não incluídas no *ladder* utilizado, em 2 dos 23 marcadores (tabela 9). Nos marcadores DYS458 e DYS385a/b, o Y-STR DYS458 foi o *locus* com o maior número de variantes, designadamente, as microvariantes 16.2, 17.2, 18.2 e 20.2, que correspondem a unidades de repetição incompletas (dois pares de bases

na unidade de repetição em vez de quatro), tendo resultados semelhantes sido descritos por Coble *et al* (Coble, Hill e Butler, 2013). O alelo 18.2 ocorreu em 8 indivíduos, tendo sido a microvariante mais comum para o marcador DYS458, o que se encontra de acordo com a base de dados www.yhrd.org.

Para o marcador DYS385a/b foi apenas identificada uma única microvariante, 12.2, previamente descrita por Terali *et al* e Dönbak *et al* (Dönbak *et al.*, 2006; Terali *et al.*, 2014).

O estudo das microvariantes alélicas é relevante não só para uma melhor compreensão da genética populacional, contribuindo também para um aumento do poder de discriminação entre indivíduos na resolução de perícias genéticas (Terali *et al.*, 2014).

Tabela 9 - Microvariantes alélicas observadas.

Loci	Alelo	Observado
DYS458	16.2	2
	17.2	2
	18.2	8
	20.2	1
DYS385a/b	12.2	1

3.4. Validação do método de ensaio

3.4.1. Estudo de concordância

3.4.1.1. Comparação com *AmpFISTR® Yfiler®*

A análise comparativa dos perfis genéticos resultantes do estudo de 66 amostras com os kits *AmpFISTR® Yfiler®* e *PowerPlex® Y23* não revelou qualquer discrepância entre eles, sendo os resultados alcançados com os dois kits concordantes entre si. Nesta amostra foram identificados 66 haplótipos diferentes, todos eles únicos, não tendo sido por isso detetadas quaisquer diferenças em termos de diversidade haplotípica entre os dois kits. O valor resultante do cálculo deste parâmetro, 1, demonstra que a presença de 6 Y-STRs adicionais no kit *PowerPlex® Y23* nem sempre implica um aumento na diversidade haplotípica, como verificado por Nuñez *et al.* (Nuñez *et al.*, 2015). Contudo, tal poderá dever-se ao pequeno número de amostras analisadas, sendo muitos os estudos que comprovam uma redução no número de haplótipos partilhados, com a inclusão de novos marcadores altamente polimórficos, bem como um poder de discriminação significativamente superior, recorrendo ao *PowerPlex® Y23* em comparação com o *AmpFISTR® Yfiler®* (Larmuseau *et al.*, 2014; Purps *et al.*, 2014; Turrina *et al.*, 2014)

3.4.1.2. Volume de reação

Neste estudo, o volume final da reação de PCR foi reduzido para 12.5 µL por uma questão de redução de custos. Contudo, foram analisadas 36 amostras num

volume final de 25 μ L e 6.25 μ L. Os perfis genéticos obtidos sob as diferentes condições foram concordantes entre si, como se denota na figura 7, exemplos de perfis obtidos para uma mesma amostra amplificada nos diversos volumes. A análise dos eletroforegramas dos três ensaios revelou que, os alelos coincidiam em todos os *loci* porém, foi nos ensaios com 25 μ L que os alelos do marcador DYS385a/b apresentavam algum desequilíbrio comparativamente aos ensaios com 12.5 μ L e 6.25 μ L. Assim, foi possível obter perfis genéticos completos e coincidentes em todos os casos analisados.

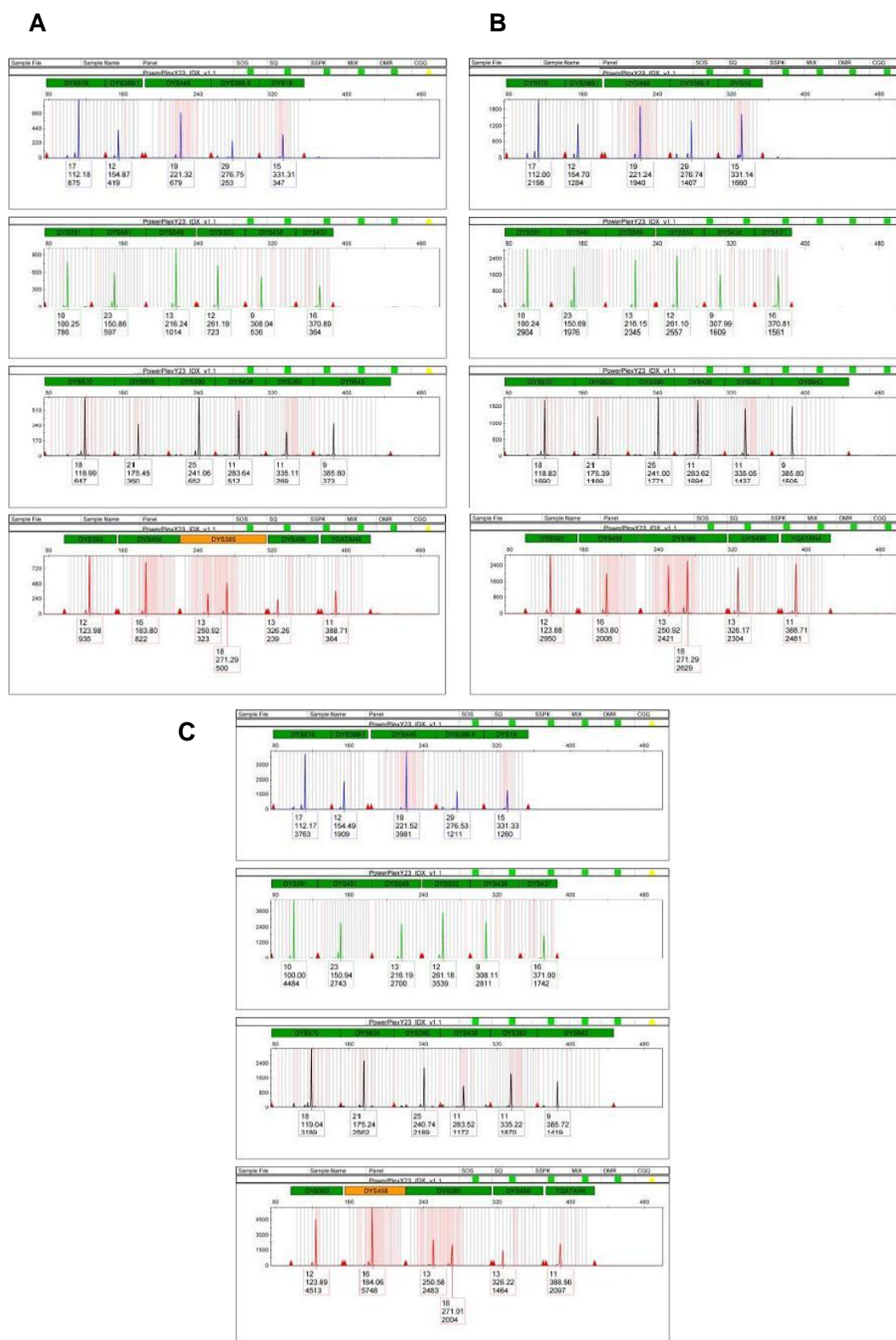


Figura 7 - Eletroforegramas da mesma amostra amplificada com o kit PowerPlex® Y23 com um volume final de 25µL (A), 12.5µL (B) e 6.25µL (C).

3.4.2. Artefactos da PCR

Os valores de *stutter* foram determinados nas 243 amostras analisadas com o *kit PowerPlex® Y23*, encontrando-se representados na tabela 10. Verifica-se que para os artefactos que correspondem a uma unidade de repetição a menos que o alelo a determinar ($n-1$), o valor de *stutter* mais elevado foi observado no marcador DYS481 (0.1904). Neste caso, em média, a dimensão do pico *stutter* é superior a 15%, o que se encontra de acordo com a literatura (Gehrig e Castella, 2013; Thompson *et al.*, 2013). Este STR por ser trinucleotídico é mais propício a *stutters* do que STRs com quatro nucleótidos por unidade de repetição (Corporation, 2012). O marcador DYS643 apresentou o menor valor de *stutter*. Relativamente aos artefactos com uma unidade de repetição a mais que o verdadeiro alelo ($n+1$), o maior e menor valor de *stutter* foi observado, respetivamente, nos marcadores DYS392 e DYS643.

Tabela 10 - Média dos valores de *stutter* para n-1 e n+1 por marcador, determinados nas 243 amostras analisadas com o *kit PowerPlex® Y23*.

<i>Locus</i>	<i>Stutter (n-1)</i>	<i>Stutter (n+1)</i>
DYS576	0.0955	0.0043
DYS389I	0.0456	0.0057
DYS448	0.0095	0.0002
DYS389II	0.1127	0.0012
DYS19	0.0564	0.0003
DYS391	0.0748	0.0036
DYS481	0.1904	0.0182
DYS549	0.0653	0.0039
DYS533	0.0582	0.0031
DYS438	0.0151	0.0009
DYS437	0.0274	0.0007
DYS570	0.0850	0.0020
DYS635	0.0586	0.0020
DYS390	0.0697	0.0017
DYS439	0.0450	0.0007
DYS392	0.0735	0.0294
DYS643	0.0034	0.0001
DYS393	0.0920	0.0118
DYS458	0.1041	0.0034
DYS385a/b	0.0804/ 0.0741	0.0017/ 0.0008
DYS456	0.1155	0.0102
YGATAH4	0.0653	0.0014

3.4.3. Balanceamento dos heterozigóticos (DYS385a/b)

Para o único marcador *multilocus* presente no kit *PowerPlex*[®] Y23, o DYS385a/b, foi determinado o desequilíbrio alélico *entre* os picos dos alelos deste *loci* heterozigótico.

Das 243 amostras analisadas com o kit *PowerPlex*[®] Y23, foram selecionados os 213 indivíduos heterozigóticos para o DYS385a/b, tendo sido obtida uma percentagem média de 82.40%, encontrando-se por isso os alelos bem equilibrados.

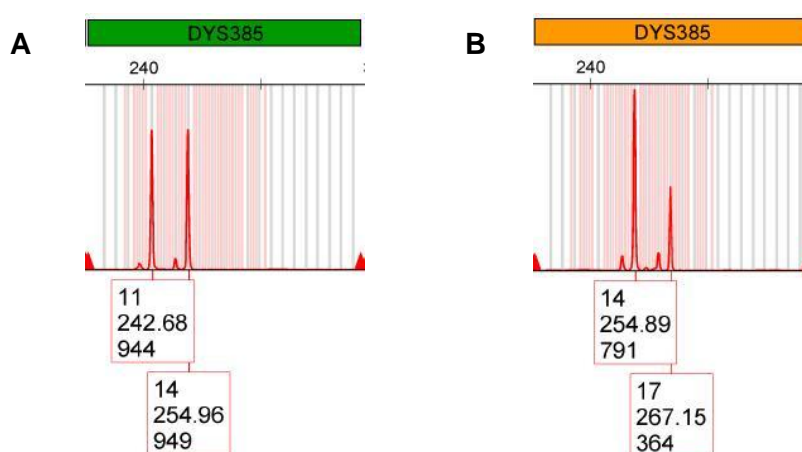


Figura 8 - Eletroforegrama exemplificativo do marcador DYS385a/b em equilíbrio alélico (A) e em desequilíbrio alélico (B).

O mesmo cálculo foi realizado para as 66 amostras tipadas com os dois *kits* utilizados neste estudo. Dos 66 perfis genéticos analisados, 55 eram heterozigóticos para o marcador DYS385a/b. Obteve-se uma percentagem média de 80.67% e 83.67% para o kit *AmpFISTR*[®] *YFiler*[®] e para o kit *PowerPlex*[®] Y23, respetivamente.

Verifica-se, em ambos os casos, a existência de equilíbrio alélico no marcador DYS385a/b, sendo este superior quando a amplificação é feita com recurso ao *kit PowerPlex® Y23*.

IV – Conclusões

O presente trabalho permitiu avaliar a adequabilidade do *kit PowerPlex® Y23* à atividade pericial desenvolvida no SGBF-S.

A genotipagem de 243 amostras de indivíduos do sexo masculino, saudáveis e não aparentados, possibilitou o estudo das características genéticas de 23 Y-STRs e de parâmetros estatísticos de interesse forense.

Os resultados obtidos revelam uma variabilidade alélica considerável por marcador genético estudado, destacando-se, na grande maioria dos casos, um alelo com frequência superior comparativamente aos outros alelos do mesmo *locus*, do qual é exemplo, o genótipo 11-14 do marcador DYS385a/b.

Em termos de diversidade génica, dois dos novos Y-STRs presentes no *kit PowerPlex® Y23*, o DYS576 e o DYS570, apresentam os valores mais elevados para este parâmetro, o que se poderá traduzir num aumento do poder de discriminação entre indivíduos, sendo tal vantajoso a nível forense.

Os elevados valores de diversidade haplotípica e de capacidade de discriminação obtidos, respetivamente 0.9998 e 0.9630, revelam a importância da utilização de um *kit* com o maior número de marcadores possíveis, reduzindo assim as possibilidades de dois indivíduos não aparentados partilharem o mesmo haplótipo.

Relativamente ao estudo comparativo entre os *kits PowerPlex® Y23* e *AmpFISTR® YFiler®*, os resultados conseguidos foram concordantes entre si e todos os parâmetros estatísticos iguais, embora fosse esperado um aumento do poder de

discriminação com o *kit PowerPlex® Y23*. Neste caso com o aumento da amostra em estudo poderia vir a observar-se uma melhoria nos parâmetros estatísticos de interesse forense.

Quanto aos estudos de validação do método realizados, nomeadamente, comparação com *AmpFISTR® YFiler®*, onde os resultados obtidos foram concordantes entre si, redução do volume final da reação de PCR, onde se alcançaram os mesmos resultados para um volume final de 6.25µL, 12.5 µL e 25 µL, estudo dos artefactos da PCR e balanceamento dos heterozigóticos, onde os valores calculados estão dentro dos valores aceitáveis, permitem concluir que o *kit PowerPlex® Y23* é aplicável à rotina laboratorial e adequado ao uso genético e forense, tendo demonstrado ser bastante sensível e robusto.

V - Referências Bibliográficas

ALBERTS, Bruce *et al.* - Química Celular e Biossíntese. Em ALBERTS, BRUCE *et al.* (Eds.) - **Biologia Molecular da Célula**. 5ª Ed ed. Porto Alegre : Artmed, 2010. p. 45–124.

ALBERTS, Bruce *et al.* - DNA, Cromossomos e Genomas. Em ALBERTS, BRUCE *et al.* (Eds.) - **Biologia Molecular da Célula**. 5ª Ed ed. Porto Alegre : Artmed, 2010. p. 195–262.

ALBERTS, Bruce *et al.* - Replicação, Reparo e Recombinação do DNA. Em ALBERTS, BRUCE *et al.* (Eds.) - **Biologia Molecular da Célula**. 5ª Ed ed. Porto Alegre : Artmed, 2010. p. 263–328.

APPLIED BIOSYSTEMS - AmpF I STR ® Yfiler ® PCR Amplification Kit. Carlsbad, CA. 4359513 (2012) 128.

BACHTROG, Doris - Y chromosome evolution: emerging insights into processes of Y chromosome degeneration. **Nat Rev Genet.** 14:2 (2014) 113–124. doi: 10.1038/nrg3366.Y.

BALLANTYNE, Kaye N. *et al.* - Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. **American journal of human genetics.** . ISSN 1537-6605. 87:3 (2010) 341–53. doi: 10.1016/j.ajhg.2010.08.006.

BALLANTYNE, Kaye N.; KAYSER, Manfred - Additional Y-STRs in Forensics: Why, Which, and When. Em SHEWALE, JAIPRAKASH G.; LIU, RAY H. (Eds.) - **Forensic DNA Analysis: Current Practices and Emerging Technologies**. Boca Raton : CRC Press, 2014. p. 221–248.

BROOKES, Clare *et al.* - Characterising stutter in forensic STR multiplexes. **Forensic science international. Genetics.** . ISSN 1878-0326. 6:1 (2012) 58–63. doi: 10.1016/j.fsigen.2011.02.001.

BUTLER, J. M. - Commonly used Short Tandem Repeat Markers and Commercial Kits. Em BUTLER, J. M. (Ed.) - **Forensic DNA Typing: Biology, Technology, and Genetics of the STR Markers**. 2^a Ed ed. USA : Elsevier, 2005. p. 85–122.

BUTLER, J. M. - Y Chromosome DNA Testing. Em BUTLER, J. M. (Ed.) - **Forensic DNA Typing: Biology, Technology, and Genetics of the STR Markers**. 2^a Ed ed. USA : Elsevier, 2005. p. 201–240.

BUTLER, J. M. - Overview and History of DNA Typing. Em BUTLER, J. M. (Ed.) - **Fundamentals of Forensic DNA Typing**. USA : Elsevier, 2010. p. 1–18.

BUTLER, J. M. - Basics of DNA Biology and Genetics. Em BUTLER, J. M. (Ed.) - **Fundamentals of Forensic DNA Typing**. USA : Elsevier, 2010. p. 19–42.

BUTLER, J. M. - Statistical Interpretation: Evaluating the Strength of Forensic DNA Evidence. Em BUTLER, J. M. (Ed.) - **Fundamentals of Forensic DNA Typing**. USA : Elsevier, 2010. p. 229–258.

BUTLER, John M. - Genetics and genomics of core short tandem repeat loci used in human identity testing. **Journal of Forensic Sciences**. . ISSN 00221198. 51:2 (2006) 253–265. doi: 10.1111/j.1556-4029.2006.00046.x.

BUTLER, John M. - Genetics and genomics of core short tandem repeat loci used in human identity testing. **Journal of forensic sciences**. . ISSN 0022-1198. 51:2 (2006) 253–65. doi: 10.1111/j.1556-4029.2006.00046.x.

CARRACEDO, Angel - **Forensic DNA Typing Protocols**. Totowa, New Jersey : Humana Press, 2005

COBLE, Michael D.; HILL, Carolyn R.; BUTLER, John M. - Haplotype data for 23 Y-chromosome markers in four U.S. population groups. **Forensic science international. Genetics**. 7:3 (2013) e66–e68.

CORPORATION, Promega - PowerPlex ® Y23 System. USA. 2012) 75.

DIEGOLI, Toni M. - Forensic typing of short tandem repeat markers on the X and Y. **Elsevier Ireland Ltd.** . ISSN 1872-4973. 2015). doi: 10.1016/j.fsigen.2015.03.013.

DÖNBAK, Lale *et al.* - Y-STR haplotypes in populations from the Eastern Mediterranean region of Turkey. **International Journal of Legal Medicine.** . ISSN 09379827. 120:6 (2006) 395–396. doi: 10.1007/s00414-005-0076-4.

FAN, Hao; CHU, Jia-You - A Brief Review of Short Tandem Repeat Mutation. **Genomics, Proteomics & Bioinformatics.** . ISSN 16720229. 5:1 (2007) 7–14. doi: 10.1016/S1672-0229(07)60009-6.

GEHRIG, Christian; CASTELLA, Vincent - Internal validation study of the PowerPlex® Y23 system. **Forensic Science International: Genetics Supplement Series.** . ISSN 18751768. 2013) 1–2. doi: 10.1016/j.fsigss.2013.10.040.

GEPPERT, Maria; EDELMANN, Jeanett; LESSIG, Ruediger - The Y-chromosomal STRs DYS481, DYS570, DYS576 and DYS643. **Legal medicine (Tokyo, Japan).** . ISSN 1873-4162. 11 Suppl 1:2009) S109–10. doi: 10.1016/j.legalmed.2009.01.063.

GILL, P. *et al.* - DNA commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. **International journal of legal medicine.** . ISSN 0937-9827. 114:6 (2001) 305–9.

GILL, P.; SPARKES, B.; BUCKLETON, J. S. - Interpretation of simple mixtures of when artefacts such as stutters are present - With special reference to multiplex STRs used by the Forensic Science Service. **Forensic Science International.** . ISSN 03790738. 95:1998) 213–224. doi: 10.1016/S0379-0738(98)00094-2.

GOODWIN, William; LINACRE, Adrian; HADI, Sibte - Lineage markers. Em GOODWIN, WILLIAM; LINACRE, ADRIAN; HADI, SIBTE (Eds.) - **An Introduction to Forensic Genetics.** England : Wiley, 2007. ISBN 978-0-470-01026-6. p. 125–136.

GUSMÃO, L. *et al.* - DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. **Forensic Science**

International. . ISSN 03790738. 160:2006) 90–101. doi: 10.1016/j.forsciint.2006.04.009.

GUSMÃO, L. *et al.* - DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. **Forensic science international.** . ISSN 0379-0738. 157:2-3 (2006) 187–97. doi: 10.1016/j.forsciint.2005.04.002.

GUSMÃO, Leonor *et al.* - Y chromosome specific polymorphisms in forensic analysis. **Legal Medicine.** . ISSN 13446223. 1:2 (1999) 55–60. doi: 10.1016/S1344-6223(99)80013-3.

HANSON, Erin K.; BALLANTYNE, Jack - Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications. **Legal Medicine.** . ISSN 1344-6223. 8:2 (2006) 110–120. doi: <http://dx.doi.org/10.1016/j.legalmed.2005.10.001>.

HU, Na *et al.* - Current developments in forensic interpretation of mixed DNA samples (Review). **Biomedical reports.** . ISSN 2049-9434. 2:3 (2014) 309–316. doi: 10.3892/br.2014.232.

HUGHES, Jennifer F.; ROZEN, Steve - Genomics and genetics of human and primate y chromosomes. **Annual review of genomics and human genetics.** . ISSN 1545-293X. 13:2012) 83–108. doi: 10.1146/annurev-genom-090711-163855.

KASHYAP, V. K. *et al.* - Deletions in the Y-derived amelogenin gene fragment in the Indian population. **BMC medical genetics.** . ISSN 1471-2350. 7:2006) 37. doi: 10.1186/1471-2350-7-37.

KAYSER, Manfred *et al.* - Characteristics and Frequency of Germline Mutations at Microsatellite Loci from the Human Y Chromosome , as Revealed by Direct Observation in Father / Son Pairs. **DNA Sequence.** 3:2000) 1580–1588.

KAYSER, Manfred; BALLANTYNE, Kaye N. - Y Chromosome in Forensic Science. Em PRIMORAC, DRAGAN; SCHANFIELD, MOSES (Eds.) - **Forensic DNA**

Applications: An Interdisciplinary Perspective. Boca Raton : CRC Press, 2014. p. 105–134.

LARMUSEAU, Maarten H. D. *et al.* - A substantially lower frequency of uninformative matches between 23 versus 17 Y-STR haplotypes in north Western Europe. **Forensic science international. Genetics.** . ISSN 1878-0326. 11:2014) 214–9. doi: 10.1016/j.fsigen.2014.04.002.

MULERO, Julio J. *et al.* - Development and validation of the AmpF ℓ STR® Yfiler™ PCR amplification kit: A male specific, single amplification 17 Y-STR multiplex system. **Journal of Forensic Sciences.** . ISSN 00221198. 51:1 (2006) 64–75. doi: 10.1111/j.1556-4029.2005.00016.x.

MULERO, Julio J.; HENNESSY, Lori K. - Next-Generation STR Genotyping Kits for Forensic Applications. Em SHEWALE, JAIPRAKASH G.; LIU, RAY H. (Eds.) - **Forensic DNA Analysis: Current Practices and Emerging Technologies.** Boca Raton : CRC Press, 2014. p. 163–180.

NAVARRO-COSTA, Paulo - Sex, rebellion and decadence: The scandalous evolutionary history of the human Y chromosome. **Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease.** . ISSN 0925-4439. 1822:12 (2012) 1851–1863. doi: <http://dx.doi.org/10.1016/j.bbadis.2012.04.010>.

NOORDAM, Michiel J. - **The human Y chromosome: a sole survivor.** [S.l.] : University of Amsterdam, 2012

NUÑEZ, Carolina *et al.* - Highly discriminatory capacity of the PowerPlex®Y23 System for the study of isolated populations. **Forensic Science International: Genetics.** . ISSN 18724973. 2015). doi: 10.1016/j.fsigen.2015.04.005.

NUSSBAUM, Robert; MCINNES, Roderick R.; WILLARD, Huntington F. - **Thompson & Thompson Genetics in Medicine.** 7. ed. Canada. 2007.

PINHEIRO, M. F. - A perícia em genética e biologia forense - criminalística biológica. Em PINHEIRO, M. F. (Ed.) - **CSI Criminal**. Porto : Edições Universidade Fernando Pessoa, 2008. p. 7–38.

PINHEIRO, M. F. - Algumas perspectivas da identificação genética. Em PINHEIRO, M. F. (Ed.) - **Genética forense: perspectivas da identificação genética**. Porto : Edições Universidade Fernando Pessoa, 2010. p. 17–78.

PINHEIRO, M. F. - Inovações em genética forense: sua contribuição na aplicação da justiça. Em PINHEIRO, M. F. (Ed.) - **Ciências forenses: ao serviço da justiça**. Lisboa : Lidel, 2013. p. 63–118.

PINHEIRO, M. Fátima - **Genética Forense Perspectivas da Identificação Genética**. 1. ed. Porto : Edições Universidade Fernando Pessoa, 2010

PINTO, Nádía; GUSMÃO, Leonor; AMORIM, António - Mutation and mutation rates at Y chromosome specific Short Tandem Repeat Polymorphisms (STRs): A reappraisal. **Forensic science international. Genetics**. . ISSN 1878-0326. 9:2014) 20–4. doi: 10.1016/j.fsigen.2013.10.008.

PLANZ, John V.; HALL, Thomas A. - Hidden Variation in Microsatellite Loci: Utility and Implications for Forensic DNA. Em SHEWALE, JAIPRAKASH G.; LIU, RAY H. (Eds.) - **Forensic DNA Analysis: Current Practices and Emerging Technologies**. Boca Raton : CRC Press, 2014. p. 199–220.

PRIMORAC, Dragan; SCHANFIELD, Moses; MARJANOVIĆ, Damir - Basic Genetics and Human Genetic Variation. Em PRIMORAC, DRAGAN; SCHANFIELD, MOSES (Eds.) - **Forensic DNA Applications: An Interdisciplinary Perspective**. Boca Raton : CRC Press, 2014. p. 3–54.

PURPS, Josephine *et al.* - A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. **Forensic science international. Genetics**. . ISSN 1878-0326. 12C:2014) 12–23. doi: 10.1016/j.fsigen.2014.04.008.

ROEWER, Lutz - Y chromosome STR typing in crime casework. **Forensic science, medicine, and pathology**. . ISSN 1556-2891. 5:2 (2009) 77–84. doi: 10.1007/s12024-009-9089-5.

ROEWER, Lutz; EPPLEN, Jorg T. - Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. **Forensic Science International**. . ISSN 03790738. 53:2 (1992) 163–171. doi: 10.1016/0379-0738(92)90193-Z.

SHAMSI, M. B.; KUMAR, K.; DADA, R. - Genetic and epigenetic factors: Role in male infertility. **Indian journal of urology: IJU: journal of the Urological Society of India**. . ISSN 1998-3824. 27:1 (2011) 110–20. doi: 10.4103/0970-1591.78436.

TERALI, K. *et al.* - Population genetics of 17 Y-STR markers in Turkish Cypriots from Cyprus. **Forensic Science International: Genetics**. . ISSN 18780326. 10:1 (2014) e1–e3. doi: 10.1016/j.fsigen.2014.01.003.

THOMPSON, Jonelle M. *et al.* - The PowerPlex® Y23 System: A New Y-STR Multiplex for Casework and Database Applications. 2012) 1–6.

THOMPSON, Jonelle M. *et al.* - Developmental validation of the PowerPlex® Y23 System: a single multiplex Y-STR analysis system for casework and database samples. **Forensic science international. Genetics**. . ISSN 1878-0326. 7:2 (2013) 240–50. doi: 10.1016/j.fsigen.2012.10.013.

THOMPSON, Jonelle M. *et al.* - Developmental validation of the PowerPlex® Y23 System: A single multiplex Y-STR analysis system for casework and database samples. **Forensic science international. Genetics**. 7:2 (2013) 240–250.

TURRINA, Stefania *et al.* - Haplotype data and mutation rates for the 23 Y-STR loci of PowerPlex® Y 23 System in a Northeast Italian population sample. **International journal of legal medicine**. . ISSN 1437-1596. 2014) 10–13. doi: 10.1007/s00414-014-1053-6.

U.S. DEPARTMENT OF ENERGY HUMAN GENOME PROJECT - **The Science Behind the Human Genome Project: Understanding the Basics** 2013. Disponível em WWW:<URL:http://web.ornl.gov/sci/techresources/Human_Genome/project/info.shtml>.

URQUHART, A. *et al.* - Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers. **International journal of legal medicine**. . ISSN 0937-9827. 107:1994) 13–20. doi: 10.1007/BF01247268.

WALSH, P. Sean; METZGER, David A.; HIGUCHI, Russell - BioTechniques 30th Anniversary Gem Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material. 54:3 (2013) 134–139.

VI – Anexos

Anexo 1 – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
h1	18	13	19	29	14	11	22	12	12	12	15	16	23	23	11	13	9	14	18	11-14	17	12	1
h2	20	13	19	29	14	10	24	12	11	12	15	19	24	24	12	13	10	13	16	11-14	15	12	1
h3	18	14	19	32	16	10	24	12	12	11	14	19	23	25	10	11	10	13	15	11-14	15	13	1
h4	20	13	20	28	16	11	21	10	11	10	15	21	22	22	12	11	12	13	17	12-12	14	12	1
h5	18	13	19	29	14	11	22	13	12	12	15	18	23	24	11	13	10	13	16	11-14	15	12	1
h6	19	13	18	28	14	11	22	14	12	12	14	17	23	24	11	13	10	14	17	11-14	17	11	1
h7	15	12	18	28	15	10	21	12	12	9	16	17	21	24	11	11	9	12	17	14-17	13	11	1
h8	18	14	19	30	14	11	22	13	13	12	15	17	23	24	12	12	10	13	18	11-14	16	12	1
h9	18	14	19	30	14	10	22	9	12	12	15	18	23	24	12	13	10	13	16	11-14	15	12	1
h10	19	13	20	28	16	11	21	10	11	10	15	22	23	22	14	11	12	13	17	12-12	14	12	1
h11	17	14	18	30	14	11	22	13	12	12	14	16	23	23	11	13	10	13	16	10-14	16	12	1
h12	17	13	19	29	14	10	22	15	13	12	15	18	23	23	12	13	10	14	18	11-14	15	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 13	18	13	19	29	14	10	22	13	12	12	15	17	23	24	13	13	10	13	16	11-14	15	12	1
H 14	18	13	19	29	14	11	22	12	12	13	16	17	23	24	13	13	10	13	17	11-14	15	12	1
H 15	15	12	20	30	15	10	22	12	11	11	13	20	20	21	13	11	13	14	18	15-16	15	11	1
H 16	17	13	19	29	14	11	23	13	12	12	15	16	23	23	13	13	10	13	17	11-15	15	12	1
H 17	19	14	20	30	13	9	27	11	11	11	14	23	21	24	10	11	12	13	17	13-14	16	12	1
H 18	19	13	19	29	15	11	20	13	13	12	15	17	25	26	11	13	10	13	17	11-14	17	12	1
H 19	19	14	19	31	14	11	21	15	12	13	15	18	23	24	12	13	10	13	16	11-15	17	12	1
H 20	18	13	19	29	14	10	22	13	12	12	15	19	23	24	11	13	10	13	19	11-14	16	12	1
H 21	19	13	18	29	14	10	22	13	12	12	15	18	23	24	12	13	10	13	16	12-14	15	12	1
H 22	16	12	21	31	17	11	23	11	11	10	15	20	21	25	11	12	14	14	16	13-17	14	10	1
H 23	19	14	20	30	13	9	26	12	11	10	14	24	22	24	10	11	12	13	17	13-14	15	12	1
H 24	18	14	20	30	13	9	28	11	10	10	14	22	20	24	10	11	12	13	17	13-14	15	12	1
H 25	21	13	19	29	14	11	22	13	12	12	15	17	23	24	12	13	8	13	18	11-13	15	12	1
H 26	18	13	19	29	14	11	23	13	12	12	15	17	23	23	13	13	10	13	17	12-14	15	11	1
H 27	18	13	19	29	14	10	21	12	13	12	15	19	23	24	12	13	10	13	17	11-14	15	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 28	18	14	19	30	14	10	22	13	12	12	15	16	23	25	12	13	10	13	17	12-14	16	11	1
H 29	18	13	19	30	13	10	23	12	12	10	14	23	21	24	12	11	11	13	15	15-18	16	11	1
H 30	18	13	19	29	14	11	22	13	12	12	15	19	23	24	12	13	10	13	17	11-14	16	12	1
H 31	17	14	20	31	13	10	22	13	12	10	14	18	24	25	12	11	12	14	17	16-17	15	11	1
H 32	18	13	19	29	15	11	22	13	13	12	15	17	23	24	12	13	11	13	17	11-13	16	12	1
H 33	18	12	19	27	15	11	22	14	11	12	14	17	24	24	10	13	10	13	16	12-14	16	12	1
H 34	19	13	20	31	13	10	24	10	13	11	14	18	22	22	12	16	12	13	16	15-17	16	10	1
H 35	18	14	19	31	14	11	21	14	12	12	15	17	23	24	13	13	10	13	17	11-14	16	12	1
H 36	17	13	19	28	14	11	22	14	12	12	15	17	23	23	12	13	10	13	17	11-14	15	12	1
H 37	17	13	19	29	14	11	23	12	13	12	14	17	25	24	13	13	10	13	17	12-14	16	11	1
H 38	19	12	19	28	14	11	22	13	11	10	15	19	23	24	12	13	10	13	17	11-14	16	11	1
H 39	19	14	20	30	14	10	26	11	12	10	14	24	21	24	10	11	12	13	19	13-14	14	12	1
H 40	19	14	20	31	14	10	22	12	11	10	16	16	21	23	11	14	11	12	16	16-17	14	10	1
H 41	16	13	20	30	14	10	25	12	11	10	14	18	20	23	11	11	9	12	18,2	12-18	15	11	1
H 42	20	13	20	31	13	10	25	12	10	10	14	21	23	24	11	11	12	13	18	17-19	17	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 43	20	12	19	28	14	11	22	12	11	10	15	20	23	24	11	13	10	13	17	11-14	15	12	1
H 44	20	13	17	29	14	12	22	13	12	12	14	18	24	24	10	13	10	13	18	11-14	16	12	1
H 45	19	13	20	30	13	10	24	12	11	10	14	18	23	24	12	11	12	12	17	16-16	14	12	1
H 46	17	13	20	30	14	10	23	13	12	12	15	17	23	24	14	13	10	13	15	13-14	15	12	1
H 47	16	14	19	31	15	10	22	14	12	9	14	17	21	23	11	13	10	13	19	12-16	16	11	1
H 48	17	14	18	30	14	11	22	11	12	12	15	17	23	24	11	13	10	13	18	11-13	16	12	1
H 49	17	12	19	28	15	11	22	14	11	9	14	17	22	26	12	11	9	12	18	16-18	14	11	1
H 50	16	12	19	28	15	10	23	14	12	9	16	16	21	24	12	11	9	12	17	10-17	12	11	1
H 51	19	13	21	30	14	10	25	12	11	10	14	18	21	23	12	11	9	12	17,2	13-15	15	11	1
H 52	19	13	19	29	14	10	22	13	12	12	15	17	23	25	12	13	9	13	16	11-13	15	12	1
H 53	20	13	19	29	14	10	22	13	12	12	15	17	23	24	12	13	10	13	16	11-14	15	12	1
H 54	18	13	20	29	14	12	22	12	10	12	15	17	23	24	12	13	10	13	18	11-15	15	12	1
H 55	20	13	19	28	14	11	24	14	12	12	15	17	23	24	11	13	10	13	17	11-14	17	12	1
H 56	16	14	20	30	13	10	21	12	12	10	14	19	22	24	12	11	12	13	15	16-17	15	12	1
H 57	18	13	20	29	13	9	27	12	12	10	14	21	21	24	10	11	12	13	19	14-14	16	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 58	19	14	19	30	14	11	24	11	12	12	15	17	23	24	12	13	10	13	20	11-14	15	12	1
H 59	18	13	19	29	15	10	22	13	12	12	15	17	23	23	12	13	11	13	17	12-14	17	12	1
H 60	18	13	19	29	14	11	22	13	12	12	15	17	24	24	14	13	10	13	16	11-14	15	12	1
H 61	17	13	19	29	15	11	22	13	12	12	15	19	23	24	11	13	10	13	15	10-14	15	12	1
H 62	18	14	20	30	13	9	27	11	11	10	14	22	21	24	10	11	13	13	18	13-14	16	12	1
H 63	18	13	19	30	14	11	21	13	12	12	15	17	24	25	11	13	10	13	17	11-14	15	11	1
H 64	17	14	19	30	16	10	22	12	12	12	15	17	23	24	11	14	10	12	18	11-14	15	12	1
H 65	16	13	20	30	13	10	22	12	12	10	14	19	24	23	12	11	12	13	15	16-18	17	11	1
H 66	20	13	20	31	13	10	25	12	11	10	14	18	24	24	12	11	12	12	18	16-16	15	12	1
H 67	20	13	19	29	14	11	22	13	13	12	15	18	23	24	12	13	11	13	17	11-14	16	12	1
H 68	15	13	21	31	13	10	23	12	12	10	14	18	22	25	12	11	13	14	16	17-17	18	12	1
H 69	20	13	20	29	14	11	22	13	12	12	15	18	23	24	12	13	10	13	17	11-14	15	11	1
H 70	17	14	19	30	14	10	22	13	12	12	14	18	23	24	11	13	10	13	16	11-15	15	12	1
H 71	18	13	19	29	14	12	23	13	11	12	14	16	23	24	12	13	10	13	17	11-17	15	12	1
H 72	18	14	19	30	13	10	28	12	11	10	14	22	22	24	9	11	11	13	18	13-14	17	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 73	14	13	20	30	12	10	24	11	11	11	14	20	22	21	12	11	13	14	18	17-17	16	11	1
H 74	16	13	19	30	14	10	23	12	12	9	14	18	21	24	12	15	10	13	17	14-16	15	12	1
H 75	16	14	18	30	14	10	21	12	12	12	15	18	23	24	13	13	9	13	17	11-14	16	11	1
H 76	17	12	19	29	15	10	23	13	12	9	16	18	21	25	11	11	9	12	16	13-18	13	11	1
H 77	17	13	19	29	14	11	23	12	11	9	15	17	22	23	12	11	11	12	17	13-16	16	11	1
H 78	20	12	21	29	14	10	22	10	11	9	14	17	22	23	11	11	9	12	17	13-13	15	11	1
H 79	16	13	20	29	14	10	23	11	11	9	15	17	22	23	10	11	11	12	15	13-16	15	11	1
H 80	19	15	17	31	14	10	22	14	12	12	14	17	23	24	12	13	10	13	17	11-16	17	11	1
H 81	16	13	21	29	14	10	23	14	12	9	15	19	23	23	11	11	10	12	17	13-16	15	12	1
H 82	19	14	20	30	13	9	27	12	11	10	14	22	21	24	10	11	12	13	18	13-15	16	12	1
H 83	19	14	19	30	14	11	22	13	12	12	15	18	23	24	12	13	10	13	18	11-14	16	12	1
H 84	16	14	20	31	17	10	27	10	11	11	14	16	21	21	12	12	13	13	16	16-18	15	11	1
H 85	18	13	20	27	16	11	22	10	12	10	15	19	22	22	11	11	12	13	17	11-12	14	12	1
H 86	17	13	21	31	16	10	23	12	11	9	14	17	21	22	11	11	8	12	15	13-15	15	12	1
H 87	18	14	20	30	13	9	27	11	11	10	14	21	21	24	10	11	12	13	17	13-15	16	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 88	19	14	19	30	14	11	22	13	12	12	15	17	23	22	13	13	10	13	17	11-14	16	12	1
H 89	18	12	20	29	15	10	25	11	10	10	15	19	19	24	12	12	12	14	18	14-15	14	11	1
H 90	15	12	23	29	15	10	22	12	10	10	16	17	22	21	11	11	12	16	17	13-17	15	11	1
H 91	17	13	19	30	14	10	22	12	13	12	15	17	23	23	12	13	10	13	19	10-15	16	12	1
H 92	17	12	19	27	16	10	22	12	9	10	16	17	22	22	11	11	12	13	18	12-14	14	11	1
H 93	18	14	19	30	14	11	24	12	12	12	16	18	23	25	12	13	10	12	17	11-14	15	12	1
H 94	19	14	20	30	13	9	26	11	11	10	14	22	21	24	10	11	12	13	17	13-14	15	12	1
H 95	19	14	20	30	13	9	27	11	11	10	14	22	21	24	10	11	12	13	19	14-14	16	12	1
H 96	18	13	20	30	14	10	25	12	11	10	14	17	20	23	11	11	9	12	18,2	14-17	15	11	1
H 97	18	13	18	29	14	11	22	15	11	12	14	16	23	24	11	13	10	13	17	11-14	16	11	1
H 98	18	14	20	31	14	9	25	13	11	10	14	22	21	23	10	11	12	13	18	12-14	18	12	1
H 99	19	14	20	31	16	10	26	11	13	10	15	18	21	21	12	12	12	14	16	16-17	13	11	1
H 100	18	13	19	30	12	10	23	12	12	12	15	17	23	24	12	13	10	13	17	11-15	16	13	1
H 101	20	13	19	30	14	10	24	11	11	10	14	18	20	22	12	11	9	12	18,2	14-18	15	11	1
H 102	17	13	19	28	14	11	22	12	12	12	15	17	23	23	12	14	11	12	17	10-15	15	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 103	18	13	20	31	14	10	24	11	10	10	14	17	20	22	12	11	9	12	16,2	14-18	15	11	1
H 104	19	13	21	28	16	10	23	10	12	10	14	19	21	23	12	11	13	13	16	12-12	14	12	1
H 105	17	13	21	27	17	10	23	10	12	10	15	19	22	23	11	11	12	13	16	12-12	14	12	1
H 106	17	15	19	31	15	11	22	12	13	12	15	17	23	24	13	13	10	13	17	11-15	15	12	1
H 107	17	13	19	28	15	11	22	14	12	12	15	18	23	24	13	13	10	13	18	11-14	15	12	1
H 108	17	14	20	30	13	9	28	11	11	10	14	22	21	23	10	11	12	13	17	13-14	14	12	1
H 109	17	13	19	29	15	11	23	13	14	12	15	18	23	24	11	13	9	13	16	11-14	15	12	1
H 110	17	13	20	30	15	11	24	12	13	11	14	18	23	25	11	11	11	13	15	11-14	17	11	1
H 111	20	14	19	31	13	10	27	13	11	9	13	17	22	25	13	14	11	11	19	15-17	15	11	1
H 112	17	14	20	30	13	10	26	11	11	10	14	24	21	24	10	11	12	13	19	13-14	14	12	1
H 113	16	12	19	28	14	10	21	13	14	12	16	18	23	25	14	13	11	14	15	11-14	15	12	1
H 114	17	14	18	30	13	9	29	11	11	10	14	22	21	24	10	11	12	13	19	13-14	16	12	1
H 115	19	13	20	32	14	10	25	13	11	10	14	18	20	23	11	11	9	12	20,2	13-16	15	12	1
H 116	17	13	19	29	14	11	22	13	12	12	15	17	23	23	12	13	10	13	18	11-14	17	11	1
H 117	19	13	20	29	14	11	22	14	12	12	15	17	25	24	12	13	9	13	18	11-14	15	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 118	20	13	21	30	13	10	22	12	12	10	14	21	23	24	12	11	12	12	15	16-19	16	11	1
H 119	16	14	19	30	15	10	22	13	12	9	14	17	21	23	11	13	10	12	18	14-14	14	11	1
H 120	16	14	20	30	15	10	22	12	11	9	14	18	21	23	11	11	10	12	17	13-15	16	11	1
H 121	19	14	19	30	14	11	22	13	12	12	15	17	23	23	12	13	10	14	17	11-14	15	12	1
H 122	18	12	21	28	15	11	24	12	11	9	16	19	23	22	11	11	8	12	17	13-16	15	11	1
H 123	16	13	24	30	15	10	23	12	9	10	16	17	21	21	12	12	12	14	17	13-16	15	11	1
H 124	16	13	21	29	17	10	24	11	11	11	14	16	21	21	12	11	13	14	16	17-18	16	13	1
H 125	19	14	19	30	13	11	23	14	12	12	15	17	24	23	11	13	11	13	18	11-14	15	12	1
H 126	19	14	19	30	14	11	22	14	12	12	15	17	23	24	12	13	10	13	16	11-15	15	12	1
H 127	18	14	21	30	14	11	23	11	11	9	15	16	21	23	11	11	11	12	15	16-16	15	11	1
H 128	17	13	19	29	14	11	23	11	11	12	15	18	23	23	11	13	10	13	18	11-13	16	12	1
H 129	16	13	19	29	14	11	22	11	12	12	15	17	23	24	12	13	10	13	17	11-15	12	12	1
H 130	18	14	19	29	14	10	22	12	12	12	15	17	24	24	13	14	10	12	17	12.2-14	15	13	1
H 131	18	12	19	28	14	11	22	13	12	13	15	16	23	24	11	12	10	13	19	12-15	17	12	1
H 132	16	13	18	28	14	11	22	13	12	12	15	17	23	23	12	13	10	13	18	11-14	16	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 133	18	13	21	30	13	10	23	12	12	10	14	21	22	24	11	10	11	13	15	15-18	15	12	1
H 134	18	13	21	31	13	10	25	13	11	10	14	18	20	23	11	11	9	12	17,2	13-17	16	11	1
H 135	17	14	19	30	14	11	22	12	12	12	15	17	23	24	12	13	10	13	17	11-14	16	11	1
H 136	16	12	21	31	17	11	23	11	11	10	15	20	21	25	11	11	14	14	16	13-17	14	10	1
H 137	15	13	21	29	14	10	25	12	11	10	16	18	22	23	11	11	12	13	16	13-14	14	11	1
H 138	19	13	21	29	14	10	22	12	12	12	15	18	23	24	12	13	10	13	17	11-15	15	12	1
H 139	19	13	20	30	16	10	23	12	12	11	14	18	23	25	11	11	10	14	15	11-14	15	13	1
H 140	16	13	21	30	14	10	22	12	11	10	14	19	20	23	11	11	12	13	16	17-19	16	11	1
H 141	19	13	19	30	15	11	22	13	12	12	15	17	23	24	12	13	10	8	17	11-13	15	12	1
H 142	18	13	19	28	14	10	22	14	12	12	15	18	24	24	11	13	10	13	17	11-15	16	12	1
H 143	17	14	20	30	13	11	27	12	11	10	14	22	21	24	10	11	12	13	17	13-14	15	12	1
H 144	18	13	19	29	14	10	22	13	13	12	14	18	25	24	12	13	10	13	16	11-14	16	12	1
H 145	19	13	20	30	13	10	24	12	11	10	14	19	25	25	13	11	12	12	17	16-17	15	12	1
H 146	15	13	21	31	15	10	24	11	11	11	14	19	21	21	11	11	13	14	18	15-15	16	11	1
H 147	15	13	20	31	15	10	22	11	11	11	14	18	21	23	11	11	13	14	15	15-16	16	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 148	19	13	20	29	15	10	22	12	12	12	15	18	23	24	13	13	10	13	18	11-15	15	12	1
H 149	17	12	19	28	15	10	22	13	12	9	16	17	21	24	11	11	9	12	16	14-17	13	11	1
H 150	19	13	19	29	13	11	22	13	12	12	14	17	23	24	12	13	10	14	17	11-14	15	12	1
H 151	16	13	18	28	14	10	23	12	12	12	15	17	23	23	11	13	10	13	17	11-14	16	13	1
H 152	17	13	21	30	15	11	21	11	9	10	16	17	21	22	12	11	11	13	19	14-14	15	11	1
H 153	18	13	19	29	14	10	22	12	13	13	15	19	23	25	11	13	10	12	17	11-14	16	12	1
H 154	18	13	19	29	14	10	24	13	12	13	15	18	23	23	13	13	9	13	16	11-14	15	12	1
H 155	18	13	20	30	14	10	24	12	11	9	15	15	22	23	11	11	10	12	18	15-16	16	10	1
H 156	17	13	19	29	14	10	24	12	12	11	15	17	25	24	12	13	10	13	16	11-14	16	12	1
H 157	17	14	20	30	13	9	25	11	11	10	14	22	20	24	10	11	12	13	20	13-13	15	12	1
H 158	21	13	19	30	14	11	22	13	13	12	15	18	23	24	12	13	11	13	17	11-14	15	12	1
H 159	18	14	19	30	14	10	22	13	12	12	15	17	23	26	11	13	10	13	16	11-15	16	12	1
H 160	14	13	21	30	16	10	25	12	12	11	14	18	21	21	13	11	12	14	18	15-16	14	11	1
H 161	20	13	20	31	14	10	24	11	10	10	14	17	20	22	13	11	9	12	18,2	14-18	15	11	1
H 162	16	13	20	30	15	10	25	11	11	10	14	18	20	23	11	11	9	12	18,2	13-18	15	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 163	18	13	19	29	14	11	22	12	11	12	15	17	24	24	11	13	11	13	17	9-14	16	11	1
H 164	18	13	18	29	14	11	23	13	12	12	15	19	23	24	12	13	9	13	19	11-14	15	12	1
H 165	18	13	20	29	17	11	26	13	13	11	15	19	20	23	12	12	13	14	15	15-15	14	11	1
H 166	16	14	21	32	15	11	28	12	11	11	14	20	21	21	12	11	14	13	16	17-17	15	12	1
H 167	16	12	19	28	16	10	24	12	11	10	14	18	22	25	11	11	13	12	15	13-16	14	11	1
H 168	15	12	19	29	15	10	23	13	12	9	15	19	21	23	12	11	9	12	16	14-17	13	11	1
H 169	16	14	20	29	14	10	22	13	12	9	14	17	21	23	11	13	10	13	19	14-17	15	10	1
H 170	18	14	21	29	15	10	23	9	12	11	15	18	21	23	12	11	11	13	16	12-12	15	11	1
H 171	17	13	19	29	14	11	23	13	12	12	15	17	23	24	12	13	10	13	17	11-15	16	12	1
H 172	19	12	20	29	13	11	23	12	12	10	14	21	23	24	11	11	12	13	15	16-16	15	11	1
H 173	17	12	19	28	17	10	24	13	12	9	16	17	22	23	12	11	9	12	16	14-18	12	11	1
H 174	17	13	19	29	14	11	23	12	13	12	15	17	23	23	12	13	10	13	16	11-14	17	12	1
H 175	17	14	20	30	13	9	26	11	11	10	14	23	21	24	10	11	12	13	17	13-14	16	12	1
H 176	16	13	18	30	15	10	27	12	11	10	15	18	21	21	12	11	11	13	18	14-17	15	11	1
H 177	19	13	19	31	14	10	22	13	12	12	15	17	24	25	11	13	11	13	17	13-13	15	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
H 178	20	13	19	29	14	11	22	12	12	12	15	17	23	24	11	13	10	13	16	12-14	16	12	1
H 179	17	12	21	29	16	10	27	11	11	12	14	18	22	21	11	11	13	14	15	14-17	15	12	1
H 180	16	12	19/ 20	29	15	11	26	12	12	10	17	18	21	22	12	11	12	13	16	13-15	16	11	1
H 181	18	13	19	30	14	10	22	12	12	12	15	16	24	24	12	13	10	13	17	12-14	15	12	1
H 182	18	13	21	28	17	10	23	10	12	10	15	19	23	23	11	11	12	13	16	12-12	14	12	1
H 183	18	13	19	29	14	11	22	13	12	12	15	17	24	24	12	13	10	13	17	12-14	17	12	1
H 184	19	13	20	29	13	9	26	11	11	10	14	23	21	24	10	11	12	13	18	13-14	16	12	1
H 185	18	13	19	29	14	10	22	12	12	13	15	17	23	23	12	13	10	13	16	11-14	16	12	1
H 186	18	14	19	30	13	11	22	12	12	12	15	17	23	23	12	13	10	13	16	11-16	15	12	1
H 187	18	14	20	30	13	9	26	11	11	10	14	20	22	24	10	11	12	13	18	13-14	16	12	1
H 188	15	13	21	30	15	11	28	11	11	11	14	19	22	21	12	11	15	13	17	17-17	15	11	1
H 189	19	12	21	28	16	10	23	12	9	10	16	17	22	23	11	11	11	14	16	14-14	15	13	1
H 190	18	13	19	29	14	10	22	12	11	12	15	17	24	24	11	13	10	13	17	12-14	15	12	1
H 191	17	13	21	30	13	10	23	12	12	10	14	22	23	24	14	11	12	13	15	16-18	16	11	1
h 192	18	13	20	29	14	11	22	13	12	12	15	19	23	24	12	13	10	13	16	11-15	15	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
h 193	17	12	21	29	15	11	22	13	9	10	16	19	21	22	13	11	12	14	16	14-15	16	12	1
h 194	19	14	18	31	14	10	22	12	12	12	14	17	23	23	13	13	11	13	17	11-14	15	11	1
h 195	15	12	23	29	16	10	23	12	10	10	16	18	21	21	11	11	12	15	16	13-16	15	11	1
h 196	17	14	21	31	15	10	24	11	11	11	14	17	21	21	12	11	13	14	16	16-16	15	11	1
h 197	18	13	19	30	15	11	22	13	12	12	14	17	23	24	11	13	9	13	16	11-15	15	12	1
h 198	20	12	19	28	14	10	23	12	12	12	15	19	23	24	12	14	10	13	15	11-14	17	12	1
h 199	18	13	19	29	15	11	22	12	12	12	15	18	23	25	12	13	10	13	18	11-14	16	13	1
h 200	19	14	19	30	14	11	20	13	12	12	15	18	24	25	11	13	10	13	17	11-14	16	12	1
h 201	16	14	21	31	17	10	25	11	11	11	14	15	21	21	12	11	13	14	16	16-17	17	11	1
h 202	17	14	21	31	15	10	25	11	12	12	14	18	22	21	12	11	12	14	17	15-16	15	12	1
h 203	18	13	19	29	14	11	22	12	12	12	15	17	23	24	12	12	10	13	17	11-14	16	12	1
h 204	19	13	20	31	14	10	27	11	12	10	14	17	21	23	11	12	12	14	17	14-18	13	11	1
h 205	18	13	20	29	15	10	25	11	13	10	14	20	21	23	12	12	12	15	17	15-16	14	11	1
h 206	17	13	20	30	14	11	22	12	12	12	15	17	23	24	12	13	11	13	19	11-14	16	12	1
h 207	19	14	19	29	14	11	22	15	12	12	15	17	23	24	12	13	10	13	17	11-14	15	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
h 208	19	12	19	28	14	11	22	12	13	12	15	16	24	24	12	13	10	13	19	11-14	17	12	1
h 209	16	13	20	30	13	10	22	12	12	10	14	21	22	24	11	11	12	13	17	17-17	16	13	1
h 210	17	13	21	31	13	10	22	12	12	10	14	19	23	23	12	11	11	13	15	16-17	17	12	1
h 211	17	13	20	31	15	11	28	11	11	11	14	20	21	21	11	11	14	13	16	16-17	15	12	1
h 212	19	13	19	30	15	11	22	14	12	12	15	18	23	25	12	13	10	13	17	11-14	16	12	1
h 213	17	13	18	29	15	11	28	13	12	10	14	19	24	22	14	11	12	12	13	12-12	15	11	1
h 214	19	13	19	29	15	11	22	12	12	12	15	17	23	24	11	13	11	13	17	11-14	15	12	1
h 215	17	13	19	29	14	11	20	14	12	12	15	20	23	24	12	13	10	13	16	11-15	16	12	1
h216	19	13	18	29	15	10	22	12	12	12	14	18	23	24	12	13	10	13	16	11-13	17	11	1
h 208	19	12	19	28	14	11	22	12	13	12	15	16	24	24	12	13	10	13	19	11-14	17	12	1
h 209	16	13	20	30	13	10	22	12	12	10	14	21	22	24	11	11	12	13	17	17-17	16	13	1
h 210	17	13	21	31	13	10	22	12	12	10	14	19	23	23	12	11	11	13	15	16-17	17	12	1
h 211	17	13	20	31	15	11	28	11	11	11	14	20	21	21	11	11	14	13	16	16-17	15	12	1
h 212	19	13	19	30	15	11	22	14	12	12	15	18	23	25	12	13	10	13	17	11-14	16	12	1
h 213	17	13	18	29	15	11	28	13	12	10	14	19	24	22	14	11	12	12	13	12-12	15	11	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na poluição Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
h 214	19	13	19	29	15	11	22	12	12	12	15	17	23	24	11	13	11	13	17	11-14	15	12	1
h 215	17	13	19	29	14	11	20	14	12	12	15	20	23	24	12	13	10	13	16	11-15	16	12	1
h 216	19	13	18	29	15	10	22	12	12	12	14	18	23	24	12	13	10	13	16	11-13	17	11	1
h 217	16	13	19	29	14	11	23	11	12	12	15	18	23	23	12	13	10	13	19	11-15	15	12	1
h 218	17	13	20	29	13	10	20	11	12	10	14	19	22	24	12	11	12	13	16	16-17	15	12	1
h 219	18	13	21	28	16	12	22	11	12	10	15	22	22	22	13	11	12	13	17	11-13	13	13	1
h 220	17	14	20	31	16	10	29	14	12	10	14	18	20	25	12	11	10	12	16,2	10-19	15	10	1
h 221	18	14	19	30	14	11	22	12	12	12	15	16	24	23	13	13	10	13	18	11-11	18	13	1
h 222	16	14	21	30	15	9	22	12	12	9	14	19	23	23	11	11	10	12	14	13-17	16	12	1
h 223	14	13	20	31	14	10	22	11	11	11	14	19	21	23	11	11	13	14	15	15-17	16	11	1
h 224	17	13	19	29	14	11	22	13	12	12	15	15	23	25	11	13	10	13	17	11-13	15	11	1
h 225	15	13	21	30	15	11	28	11	11	11	14	19	21	21	11	11	13	13	17	17-17	15	12	1
h 226	17	13	19	29	14	10	23	14	13	11	15	18	24	24	13	13	9	13	18	11-14	16	12	1
h 227	18	13	21	30	17	10	26	12	11	11	14	14	22	22	12	11	12	13	16	17-17	15	10	1
h 228	19	13	19	29	14	10	22	12	12	12	15	17	25	24	11	13	10	13	17	9-14	16	12	1

Anexo 1 (continuação) – Haplótipos de STRs do cromossoma Y na população Sul de Portugal

Haplótipo	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385a/b	DYS456	YGATAH4	N
h 229	17	14	20	30	14	10	22	14	11	9	15	19	25	24	12	11	10	12	17	14-15	17	12	1
h 230	19	13	20	30	14	10	24	11	10	10	14	18	20	22	12	11	9	12	18,2	14-18	15	11	1
h 231	17	12	20	28	14	10	26	12	11	10	16	19	22	22	11	11	12	13	15	13-14	14	11	1
h 232	17	13	19	29	14	11	22	13	12	12	15	17	23	24	12	13	10	13	16	11-14	16	11	2
h 233	17	12	20	29	13	10	26	12	11	10	14	18	21	24	12	11	9	12	18,2	12-17	15	10	2
h 234	17	13	20	29	13	9	27	12	11	10	14	22	21	24	10	11	12	13	18	13-14	16	12	2
h 235	18	12	19	29	15	11	24	12	12	12	15	18	23	23	13	13	10	13	17	11-14	16	13	2
h 236	18	13	20	30	15	11	22	13	12	12	15	18	23	24	12	13	10	14	17	11-14	15	12	2
h 237	18	14	20	30	13	9	28	11	11	10	14	23	21	24	10	11	12	13	18	13-14	16	11	2

N – número de indivíduos observados por cada haplótipo.

Anexo 2 – Frequências alélicas dos marcadores genéticos do *kit PowerPlex® Y23*.

Locus	Alelos	Frequência
DYS576	14	0.0123
	15	0.0453
	16	0.1276
	17	0.2469
	18	0.2922
	19	0.1975
	20	0.0700
	21	0.0082
DYS389I	12	0.1399
	13	0.5844
	14	0.2675
	15	0.0082
DYS448	17	0.0082
	18	0.0617
	18,19	0.0041
	19	0.4527
	20	0.3086
	21	0.1523
	23	0.0082
	24	0.0041
DYS389II	27	0.0165
	28	0.1152
	29	0.3704
	30	0.3498
	31	0.1358
	32	0.0123
DYS19	12	0.0082
	13	0.1893
	14	0.4815
	15	0.2099
	16	0.0700
	17	0.0412
DYS391	9	0.0864
	10	0.5021
	11	0.3951
	12	0.0165

Anexo 2 (continuação) - Frequências alélicas dos marcadores genéticos do *kit* PowerPlex® Y23.

Locus	Alelos	Frequência
DYS481	20	0.0165
	21	0.0453
	22	0.4403
	23	0.1646
	24	0.1029
	25	0.0658
	26	0.0576
	27	0.0576
	28	0.0412
	29	0.0082
DYS549	9	0.0082
	10	0.0370
	11	0.2016
	12	0.3827
	13	0.2716
	14	0.0823
	15	0.0165
DYS533	9	0.0206
	10	0.0370
	11	0.3210
	12	0.5309
	13	0.0823
	14	0.0082
DYS438	8	0.0041
	9	0.0947
	10	0.3374
	11	0.0988
	12	0.4403
	13	0.0247
DYS437	13	0.0082
	14	0.4198
	15	0.4897
	16	0.0782
	17	0.0041
DYS570	14	0.0041
	15	0.0123
	16	0.0617
	17	0.3416
	18	0.2593
	19	0.1440
	20	0.0412
	21	0.0329
	22	0.0658
	23	0.0247
	24	0.0123

Anexo 2 (continuação) - Frequências alélicas dos marcadores genéticos do *kit* PowerPlex® Y23.

Locus	Alelos	Frequência
DYS635	19	0.0041
	20	0.0617
	21	0.2469
	22	0.1276
	23	0.4403
	24	0.0864
	25	0.0329
DYS390	21	0.0782
	22	0.0782
	23	0.2469
	24	0.4897
	25	0.0947
	26	0.0123
DYS439	9	0.0041
	10	0.1111
	11	0.3169
	12	0.4403
	13	0.1029
	14	0.0247
DYS392	10	0.0041
	11	0.4568
	12	0.0453
	13	0.4609
	14	0.0247
	15	0.0041
	16	0.0041
DYS643	8	0.0123
	9	0.1193
	10	0.4362
	11	0.1070
	12	0.2387
	13	0.0658
	14	0.0165
	15	0.0041
DYS393	8	0.0041
	11	0.0041
	12	0.1811
	13	0.6708
	14	0.1276
	15	0.0082
	16	0.0041

Anexo 2 (continuação) - Frequências alélicas dos marcadores genéticos do *kit* PowerPlex® Y23.

Locus	Alelos	Frequência
DYS458	13	0.0041
	14	0.0041
	15	0.0988
	16	0.2305
	16.2	0.0082
	17	0.3663
	17.2	0.0082
	18	0.1687
	18.2	0.0329
	19	0.0658
	20	0.0082
	20.2	0.0041
DYS456	12	0.0123
	13	0.0288
	14	0.0947
	15	0.4486
	16	0.3169
	17	0.0864
	18	0.0123
YGATAH4	10	0.0412
	11	0.3498
	12	0.5556
	13	0.0535

Anexo 3 - Frequência genotípica do marcador DYS385a/b.

Marcador	Genótipos	Frequência
DYS385a/b	9,14	0.0082
	10,14	0.0082
	10,15	0.0082
	10,17	0.0041
	10,19	0.0041
	11,11	0.0041
	11,12	0.0041
	11,13	0.0370
	11,14	0.2881
	11,15	0.0700
	11,16	0.0082
	11,17	0.0041
	12,12	0.0288
	12,14	0.0535
	12.2,14	0.0041
	12,15	0.0041
	12,16	0.0041
	12,17	0.0082
	12,18	0.0041
	13,13	0.0123
	13,14	0.0864
	13,15	0.0247
	13,16	0.0329
	13,17	0.0206
	13,18	0.0082
	14,14	0.0206
	14,15	0.0123
	14,16	0.0041
	14,17	0.0288
	14,18	0.0247
	15,15	0.0082
	15,16	0.0247
	15,17	0.0123
	15,18	0.0082
	16,16	0.0206
	16,17	0.0370
	16,18	0.0165
	16,19	0.0041
	17,17	0.0288
	17,18	0.0041
	17,19	0.0082

Anexo 4 – Artigo publicado



Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS



Analysis of Y chromosomal STRs in South of Portugal population with PowerPlex® Y23

M. Cruz^a, T. Ribeiro^b, M.J. Porto^{b,*}, J. Costa Santos^a, M. Carvalho^b

^a Faculty of Medicine, University of Lisboa, Lisboa, Portugal

^b Forensics Genetic Service, South Delegation, National Institute of Legal Medicine and Forensic Sciences, I.P., Lisboa, Portugal

ARTICLE INFO

Article history:

Received 27 August 2015

Accepted 17 September 2015

Available online xxx

Keywords:

Y-STRs

PowerPlex® Y23

Population genetics

ABSTRACT

Short tandem repeats (STRs) became a vital tool in forensic genetics due to its high level of polymorphism and abundance in the human genome. Y-chromosome short tandem repeats (Y-STRs) specifically, play an important role in sexual assault cases, where the female component is in greater proportion than the male component, in complex paternity testing of male children and in the process of identification in cases of missing persons or disaster victim identification, involving profiles of male relatives. The PowerPlex® Y23 System is a five-dye Y-STR multiplex that allows simultaneous genotyping of 23 Y-STR loci (DYS576, DYS389I, DYS448, DYS389II, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458, DYS385a/b, DYS456 and Y-GATA-H4). In this study, samples of 243 unrelated male individuals living in the south of Portugal were amplified in a GeneAmp® 9700 Thermal Cycler. PCR products were separated and detected on an ABI PRISM® 3130 xl and the electrophoresis results analyzed with GeneMapper® ID software v1.2. Allele frequencies and haplotype frequencies were estimated and compared with frequencies from other European population. The results show that the study of Y-STRs using PowerPlex® Y23 kit, provide a powerful discrimination tool for routine forensic applications.

© 2015 Published by Elsevier Ireland Ltd.

1. Introduction

Currently, STRs are the markers of choice in forensic investigations due to their significant variability established through a great number of loci. Y chromosome STRs are widely used in forensic genetics for lineage and male sex identification, kinship investigations and in population genetics [1].

The PowerPlex® Y23 System allows co-amplification of seventeen Y-STR loci previously available in Y-STR kits (DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and Y-GATA-H4) and six new highly discriminating Y-STR loci (DYS481, DYS533, DYS549, DYS570, DYS576 and DYS643). The high gene diversities of these six new loci contribute to an additional distinction between unrelated male individuals [2].

In this work, the study of allelic and haplotype frequencies of these 23 Y chromosome markers, was carried out in the south of Portugal population and forensic statistical parameters calculated in order to evaluate their applicability to routine forensic cases.

2. Material and methods

Samples of 243 unrelated male individuals living in the south of Portugal were amplified in a 12.5 µL reaction volume using the PowerPlex® Y23 System (PPY23). Of this samples, 36 were also amplified in a final volume of 25 µL, according to the manufacturer's recommendations [3], and a lower volume of 6.25 µL in a GeneAmp® 9700 Thermal Cycler. PCR products were separated and detected on an ABI PRISM® 3130 xl and the electrophoresis results analyzed with GeneMapper® ID software v1.2.

Allele and haplotype frequencies were determined by gene/haplotype counting. Haplotype diversity (HD) and gene diversity (GD) were calculated using the formula $n(1 - \sum p_i^2)/(n - 1)$, where for HD n represents the number of haplotypes and p_i the frequency of the i th haplotype, and for GD n is the sample size and p_i is the allelic frequency. The values of HD and GD obtained for this population were compared with the values achieved for other European populations.

Heterozygote balance was measured for DYS385a/b by dividing the minimum allele peak height by the maximum allele peak height. Stutter values were calculated by dividing the stutter peak height by the allele peak height. The peaks considered where one repeat smaller or larger than the true allele [4].

* Corresponding author.

E-mail address: mariajoao.porto@dcinml.mj.pt (M.J. Porto).

Table 1

Allele frequencies and gene diversity of 21 Y-STR loci in 243 unrelated male individuals living in the south of Portugal (N=243).

Alleles	DYS576	DYS389 I	DYS448 ^b	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS456	YGATAH4
8										0.0041							0.0123	0.0041			
9						0.0864		0.0082	0.0206	0.0947					0.0041		0.1193	0.0041			
10						0.5021		0.0370	0.0370	0.3374					0.1111	0.0041	0.4362	0.1811			0.0412
11						0.3951		0.2016	0.3210	0.0988					0.3169	0.4568	0.1070	0.6708			0.3498
12		0.1399			0.0082	0.0165		0.3827	0.5309	0.4403					0.4403	0.0453	0.2387	0.1276		0.0123	0.5556
13		0.5844			0.1893			0.2716	0.0823	0.0247	0.0082				0.1029	0.4609	0.0658	0.0082	0.0041	0.0288	0.0535
14	0.0123	0.2675			0.4815			0.0823	0.0082		0.4198	0.0041			0.0247	0.0247	0.0165	0.0041	0.0041	0.0947	
15	0.0453	0.0082			0.2099			0.0165			0.4897	0.0123				0.0041	0.0041	0.0041	0.0988	0.4486	
16	0.1276				0.0700						0.0782	0.0617				0.0041		0.0041	0.2305	0.3169	
16.2																			0.0082		
17	0.2469		0.0082		0.0412						0.0041	0.3416							0.3663	0.0864	
17.2																			0.0082		
18	0.2922		0.0617									0.2593							0.1687	0.0123	
18.2																			0.0329		
19	0.1975		0.4527								0.1440	0.0041							0.0658		
20	0.0700		0.3086				0.0165				0.0412	0.0617							0.0082		
20.2																			0.0041		
21	0.0082		0.1523				0.0453					0.0329	0.2469	0.0782							
22			0.0082				0.4403					0.0658	0.1276	0.0782							
23			0.0041				0.1646					0.0247	0.4403	0.2469							
24			0.0082				0.1029					0.0123	0.0864	0.4897							
25							0.0658						0.0329	0.0947							
26							0.0576							0.0123							
27				0.0165			0.0576														
28				0.1152			0.0412														
29				0.3704			0.0082														
30				0.3498																	
31				0.1358																	
32				0.0123																	
GD ^a	0.7912	0.5673	0.6700	0.7083	0.6816	0.5841	0.7534	0.7306	0.6065	0.6693	0.5778	0.7835	0.7165	0.6779	0.6821	0.5762	0.7223	0.5009	0.7689	0.6808	0.5644

^a Gene diversity.

^b At DYS448 a locus duplication was observed, 18–19, with a allele frequency of 0.0041.

Table 2Genotype frequency of locus DYS385a/b (*N* = 243).

Locus	Genotypes	Frequency	Genotypes	Frequency
DYS385a/b	9–14	0.0082	13–15	0.0247
	10–14	0.0082	13–16	0.0329
	10–15	0.0082	13–17	0.0206
	10–17	0.0041	13–18	0.0082
	10–19	0.0041	14–14	0.0206
	11–11	0.0041	14–15	0.0123
	11–12	0.0041	14–16	0.0041
	11–13	0.0370	14–17	0.0288
	11–14	0.2881	14–18	0.0247
	11–15	0.0700	15–15	0.0082
	11–16	0.0082	15–16	0.0247
	11–17	0.0041	15–17	0.0123
	12–12	0.0288	15–18	0.0082
	12–14	0.0535	16–16	0.0206
	12–2–14	0.0041	16–17	0.0370
	12–15	0.0041	16–18	0.0165
	12–16	0.0041	16–19	0.0041
	12–17	0.0082	17–17	0.0288
	12–18	0.0041	17–18	0.0041
	13–13	0.0123	17–19	0.0082
	13–14	0.0864		

3. Results and discussion

A total of 237 different haplotypes were found, of which 231 were unique and 6 haplotypes were observed in two individuals each. Locus duplication was observed at the DYS448 locus. Intermediate alleles 16.2, 17.2, 18.2 and 20.2 were observed at DYS458 and 12.2 at DYS385a/b. Similar results, for intermediate alleles, have been reported at these loci for a Turkish Cypriot population [5].

The calculated gene diversity values ranged from 0.5009 to 0.7912 for the DYS393 and DYS576 loci, respectively. Two of the new loci present in the PPY23 kit, DYS576 and DYS570, were among the markers with the highest GD (Table 1). For the marker DYS385a/b, 41 different genotypes were observed standing out the genotype 11–14 with a frequency of 0.2881 (Table 2).

The overall haplotype diversity was determined as 0.9998 and compared with a Northeast Italian population ($HD = 0.999994$) [6] and a set of European populations ($HD = 0.999992$) [7], being within the expected values.

The profiles obtained with the volumes 25, 12.5 and 6.25 μ L were consistent with each other, and lowering the reaction volume has a great economic impact.

For DYS385a/b the average peak height ratio was 82.40%, greater than 70%, indicating that the peaks are well balanced.

The highest stutter value was observed in the marker DYS481 (0.1904), which was expected due to be a 3 bp repeat. In turn, the lowest stutter value was observed in the marker DYS643 (0.0034).

4. Conclusions

The values of GD and HD are consistent with the ones obtained for other European population. The results from the statistical parameters calculation, together with the method validation studies, namely PCR artefacts, prove that the 23 genetic markers are a powerful discrimination tool suitable for forensic applications.

Conflict of interest

None.

References

- [1] C. Nuñez, M. Baeta, M. Fernández, M. Zarrabeitia, B. Martinez-Jarreta, M.M.d. Pancorbo, Highly discriminatory capacity of the PowerPlex® Y23 System for the study of isolated populations, *Forensic Sci. Int. Genet.* 17 (2015) 104–107 Elsevier Ireland Ltd..
- [2] J.M. Thompson, M.M. Ewing, W.E. Frank, J.J. Pogemiller, C.a. Nolde, D.J. Koehler, et al., Developmental validation of the PowerPlex® Y23 System: a single multiplex Y-STR analysis system for casework and database samples, *Forensic Sci. Int. Genet.* 7 (February (2)) (2013) 240–250 Elsevier Ireland Ltd..
- [3] Promega Technical Manual PowerPlex® Y23 System. USA: Promega Corporation; 2012. p. 75.
- [4] C. Gehrig, V. Castella, Internal validation study of the PowerPlex® Y23 System, *Forensic Sci. Int. Genet. Suppl. Ser.* (November) (2013) 1–2 Elsevier Ireland Ltd..
- [5] K. Kerali, T. Zorlu, O. Bulbul, C. Gurkan, Population genetics of 17 Y-STR markers in Turkish Cypriots from Cyprus, *Forensic Sci. Int. Genet.* 10 (1) (2014) e1–e3 Elsevier Ireland Ltd..
- [6] S. Turrina, S. Caratti, M. Ferriani, D. De Leo, Haplotype data and mutation rates for the 23 Y-STR loci of PowerPlex® Y23 System in a Northeast Italian population sample, *Int. J. Legal Med.* (August (7)) (2014) 10–13.
- [7] J. Purps, S. Siegert, S. Willuweit, M. Nagy, C. Alves, R. Salazar, et al., A global analysis of Y-chromosomal haplotype diversity for 23 STR loci, *Forensic Sci. Int. Genet.* 12 (2014) 12–23.